The Dissertation Committee for Alex Jemyung Lee
certifies that this is the approved version of the following dissertation:

Real-space electronic structure methods for modeling nanowires and atomic force microscopy imaging

Committee:

James R. Chelikowsky, Supervisor

Alexander A. Demkov

Gyeong S. Hwang

Brian A. Korgel

Isaac C. Sanchez

Real-space electronic structure methods for modeling nanowires and atomic force microscopy imaging

by

Alex Jemyung Lee, B.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2016

For my loving parents

Acknowledgments

There are numerous people I must thank who have strongly contributed to my life as a graduate student.

First and foremost, I thank my adviser Dr. Chelikowsky for all of his guidance, support, and patience in working with me throughout the years. I will always remember his stories and good humor. I also thank Robert Hoelscher for his assistance and ability to make the lives of everyone in the group much easier.

Thanks goes out to my colleagues in the lab from whom I have learned much and shared many experiences: Scotty Bobbitt, Ben Garrett, Charles Lena, Grady Schofield, Masahiro Sakurai, and Jaime Souto-Casares. A few individuals deserve special mention. First, I thank Dr. Minjung Kim for helping me gain a foothold in the group and for always being willing to answer my many questions. I thank Dr. Tzu-Liang Chan, who I have unfortunately never been able to meet in person, but who provided much helpful advice and showed openness and patience in doing multiple long-distance collaborations. I also thank Dr. Yuki Sakai for his invaluable assistance in pushing forward the atomic force microscopy simulation projects.

Thanks goes out to my family, my parents, my sister, and her husband John, for being never-ending sources of love and support.

I acknowledge support from the U.S. Department of Energy (DoE) and the Welch Foundation. Support on algorithms was provided by the Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research and Basic Energy Sciences. Computational resources were provided by the National Energy Research Scientific Computing Center (NERSC) and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation, at the Texas Advanced Computing Center (TACC). I also thank the friends and family of Harry P. Whitworth for their continued encouragement and financial support regarding their participation in the Cockrell Graduate Fellowship.

Real-space electronic structure methods for modeling nanowires and atomic force microscopy imaging

Publication No. _____

Alex Jemyung Lee, Ph.D.
The University of Texas at Austin, 2016

Supervisor: James R. Chelikowsky

Our group develops a pseudopotential-based electronic structure code constructed within *ab initio* density functional theory to study various problems in chemical engineering and condensed matter physics. One of the code's unique features is that it performs calculations on a real-space grid without using an explicit basis. This makes it particularly well-suited for examining localized systems with confined dimensionalities, such as nanostructures. In the dissertation we apply our code to the study of two main topics: germanium nanowires and atomic force microscopy simulations.

First we examine how the electronic properties of germanium nanowires are affected by mechanical strain. We find that applying strain can drastically influence the transport properties of nanowires by inducing band crossings that change the nature of the band gap from direct to indirect, hampering carrier mobilities. In another project we take advantage of the real-space formalism

for charged systems to devise a computationally efficient method to calculate accurate doping binding energies for nanowires. We demonstrate the method on phosphorus-doped germanium nanowires.

The second focus of the dissertation is atomic force microscopy simulations. Atomic force microscopy is a powerful probe-based imaging technique that can be used to visualize and characterize chemical phenomena. However, the interpretation of experimental images is not well-understood. We develop a theoretical simulation method in order to better understand the fundamental physics behind the imaging mechanism. In one study, we clarify the mechanism for imaging hydrogen bonds. Experimental findings on certain organic oligomers have reported striking images showing what appear to be direct visualizations of intermolecular bonding. We apply our simulation technique to show that tip tilting is responsible for resolving these apparent hydrogen bonds. In another study, we examine the phenomenon of contrast inversion. We find that the key factor responsible for contrast inversion is the chemical reactivity of the tip. Theoretical imaging simulations such as these can be used to guide the experimental acquisition of images and to help characterize results.

Table of Contents

List of Tables

List of Figures

Chapter 1

Dissertation overview

Computer simulations based on electronic structure theory are a useful tool for nanoscience research. Simulations can be used to study and design materials with desirable properties that take advantage of the nanoscale regime. Simulations can also model experimental processes in order to guide experimental technique and aid the interpretation of results. A common thread in the dissertation is to discuss not only the results obtained from models but also to emphasize the construction of the models themselves. Nearly all models rely on approximations; different models focus on different aspects of the physical problem and have different strengths and weaknesses based on the context of their application. A comprehensive understanding of the methodologies can allow researchers to select the most appropriate model for a specific application in order to achieve the best results.

The opening chapters provide a theoretical overview of the science and computational tools used in the dissertation. Chapter 2 introduces the electronic structure problem, density functional theory, and the PARSEC code, the main computational tool used in our group's research. Chapter 3 shows how the code's unique formalism makes it well-suited to study nanostructures,

particularly with its option for partially periodic boundary conditions.

Chapters 4 and 5 present examples of nanostructures projects that can be carried out with PARSEC. The first project examines how axial strain affects the electronic and transport properties of germanium nanowires. This chapter was adapted from a published paper by A. J. Lee, M. Kim, C. Lena, and J. R. Chelikowsky. Phys. Rev. B., 86:115331, 2012 [72]. Author contributions are as follows: A. J. Lee headed the project, ran the primary calculations, and wrote the paper; M. Kim offered technical advice; C. Lena assisted with data analysis; and J. R. Chelikowsky oversaw the project.

Chapter 5 develops a method to calculate dopant binding energies in nanowires. Understanding dopant properties is essential for semiconductors research since the dopant is primarily responsible for giving the semiconductor its desirable properties. We demonstrate the method on phosphorus-doped germanium nanowires. This chapter was adapted from a published paper by A. J. Lee, T-L. Chan, and J. R. Chelikowsky. Phys. Rev. B., 89: 075419, 2014 [71]. Author contributions are as follows: A. J. Lee headed the project, ran the primary calculations, and wrote the paper; T-L. Chan offered technical advice and made significant edits to the paper; and J. R. Chelikowsky oversaw the project.

The remainder of the dissertation covers atomic force microscopy (AFM) simulations. Atomic force microscopy is a powerful probe-based imaging technique used in surface science and nanotechnology research. Imaging can be used to characterize the detailed atomic structure of samples and aid the inter-

pretation of complex chemical features such as bond order and intermolecular bonding. Chapter 6 contains an overview of AFM theory and describes various models developed to simulate AFM images. The different models rely on different approximations and focus on different contributions to the tip-sample interaction that governs the imaging contrast mechanism. By studying and applying these models, we can identify the ideal method to simulate an AFM image for a target system.

The final two chapters present examples of research problems that can be studied with AFM simulations. Chapter 7 details the imaging mechanism behind the resolution of apparent hydrogen bonding seen in some experimental AFM images. This chapter was adapted from a published paper by A. J. Lee, Y. Sakai, M. Kim, and J. R. Chelikowsky. App. Phys. Lett., 108:193102, 2016 [73]. Author contributions are as follows: A. J. Lee headed the project, ran the primary calculations, and wrote the paper; Y. Sakai offered technical advice and assisted with running calculations; M. Kim performed foundational work in the early stages of the project; and J. Chelikowsky oversaw the project. Chapter 8 examines the phenomenon of contrast inversion seen in some experimental AFM images and investigates factors that could be responsible for the inversion, such as the chemical reactivity of the tip as well as electrostatic forces.

Chapter 2

Theoretical overview

2.1 Introduction

The main computational tool our research group uses for electronic structure calculations is the so-called PARSEC code. PARSEC is a pseudopotential-based code constructed within *ab initio* density functional theory (DFT) and uses finite difference methods to carry out calculations on a real space grid without the use of an explicit basis. The code's foundations were laid out in 1994 by James Chelikowsky, Norm Troullier, and Yousef Saad when the authors proposed a method to perform electronic structure calculations in real space [26, 69]. This approach differed from popular methods that expanded wave functions on a plane wave basis and carried out calculations in momentum or k-space. While plane wave codes are still commonly used, we describe an alternative approach using a real-space grid. The real space formalism offers various computational benefits and can mitigate some of the issues involved with using plane waves without any loss of accuracy. In this chapter we provide an overview of the theoretical underpinnings of density functional theory and the PARSEC code, highlighting some of the code's unique features and showing how these features can be applied to solve various research problems.

2.2 The Schrödinger equation and density functional theory

The fundamental physical equation at the heart of *ab initio* electronic structure calculations is the Schrödinger equation, presented here in the non-relativistic, time-independent form:

$$\hat{H}\psi = E\psi, \tag{2.1}$$

where $\hat{H}$ is a Hamiltonian operator, ψ is a wave function representing a stationary quantum state, and E is the energy eigenvalue of that state. For an atomic system composed of electrons and nuclei, the Hamiltonian can be written as:

$$\begin{aligned}\hat{H} = &-\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 + \sum_{i,I}\frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &-\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|},\end{aligned} \tag{2.2}$$

where $\hbar$ is Planck's constant h divided by 2π, e is the charge of the electron, Z_I are the charges of the nuclei, $\mathbf{r}_i$ and $\mathbf{R}_I$ refer to the positions of the electrons and nuclei, and m_e and M_I refer to the masses of the electrons and nuclei. From left to right, the five terms in the Hamiltonian represent the kinetic energies of the electrons, the nuclei-electron interactions, the electron-electron interactions, the kinetic energies of the nuclei, and the nuclei-nuclei interactions. A factor of $\frac{1}{2}$ is included where necessary to avoid double-counting the interactions.

The Hamiltonian can be greatly simplified by applying the Born-Oppenheimer approximation [10], which is typically done for most condensed matter calculations. The main premise of the approximation is that we can decouple the motions of nuclei and electrons adiabatically. Because nuclear masses are so much larger than electron masses, we can assume that electron velocities are essentially infinite compared to those of the much heavier nuclei. Electrons are then viewed as responding instantaneously to changes in the positions of the nuclei, which are regarded as fixed in comparison. We can then eliminate the nuclear kinetic energy term from the Hamiltonian. The term for the internuclei interactions also becomes a constant and is often subtracted out to simplify calculations. The new Hamiltonian becomes:

$$\begin{aligned} \hat{H}\psi &= -\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 + \sum_{i,I}\frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &= \hat{T} + \hat{V}_{ion} + \hat{V}_{int}, \end{aligned} \tag{2.3}$$

where $\hat{T}$ represents the kinetic energy operator for the electrons, $\hat{V}_{ion}$ represents the potential of the nuclei acting on the electrons, and $\hat{V}_{int}$ represents electron-electron repulsions. This Hamiltonian is the basis for the Hartree approximation method to solve the Schrödinger equation. This method turns the complex many-body problem into a simplified one-electron picture by writing the wave function in terms of orbitals that represent discrete energy states for the system. The Hartree-Fock approximation improves upon the Hartree method by accounting for electron exchange.

Even with the Hartree-Fock approximation, the Schrödinger equation

is difficult to solve exactly for many-electron systems. Hartree-Fock is a "mean field" approximation, which means that complex many-body interactions are mapped onto a single-particle picture. The potential for the system is defined to act as the average response for all of the electrons in the system, hence the term "mean field." The consequence of this simplification is that it cannot account for all of the so-called electron correlation energies, so the energies calculated by Hartree-Fock are always too high. Various methods to recover more of the correlation energy exist but are generally computationally expensive.

Density functional theory is a popular and in some ways controversial method that gets around some of the issues in the Hartree-Fock method. The foundations for modern density functional theory were laid out in 1964 with the establishment of the Hohenberg-Kohn theorems [55]. The theorems gave rise to the Kohn-Sham equations, which applied the theorems to a practical calculation method for the electronic structure problem [65]. Improvements to the accuracy and implementation of the method, including work done by Becke [64], along with sharp rises in computational power enabled widespread application of density functional theory to materials research. For work in the development of density functional theory, Walter Kohn was awarded the Nobel Prize in Chemistry alongside John Pople in 1998.

The Hohenberg-Kohn theorems state that the Hamiltonian for an electronic system, and thus the system's energy, can be determined exactly by a single function, which is the ground-state electron density. Furthermore, this ground-state density can be found by minimizing the total energy of some as-

sociated density functional. These theorems are extremely powerful because they prove the existence of an exact solution to the many electron problem in terms of a single function. However, the theorems do not provide a means to actually find the ground-state density since they give no indication on what the density functional actually looks like or even how to go about constructing it. A need for a practical calculation method quickly led to development of the Kohn-Sham equations in 1965 [65].

The Kohn-Sham equation can be seen as the density functional theory analogue to the Schrödinger equation and is written:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{eff}(\rho)\right)\phi_i(\rho) = \epsilon_i\phi_i(\rho), \tag{2.4}$$

where V_{eff} is the effective Kohn-Sham potential and ϕ_i are Kohn-Sham orbitals with corresponding eigenvalue energies ϵ_i. The Kohn-Sham orbitals are related to the electron density $\rho(\mathbf{r})$ by:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2, \tag{2.5}$$

where N is the number of occupied states.

The effective Kohn-Sham potential is split into multiple terms:

$$V_{eff}(\rho) = V_{ion}(\rho) + V_H(\rho) + V_{xc}(\rho), \tag{2.6}$$

where V_{ion} is the ionic potential encompassing electron-nuclei interactions, V_H is the Hartree potential representing electron-electron repulsions, and V_{xc} is the exchange-correlation potential. We will now take a closer look at each term in the Kohn-Sham Hamiltonian and describe how PARSEC deals with each of them.

2.3 Solving the Kohn-Sham equation in real-space

One of the unique aspects of PARSEC is that it calculates the electronic structure problem on a real-space grid without the use of an explicit basis. The "basis" is instead the number of grid points used in the calculation, which controls convergence. The real-space approach can have various technical advantages. One of the most important ones is computational efficiency. The real-space method typically produces Hamiltonian matrices that are sparse and therefore quick to diagonalize. While the matrices tend to be large since they scale with the number of grid points needed, the entire matrix does not need to be stored explicitly. Essential terms can be preserved by writing the matrix in "stencil" form, resulting in huge memory savings. Figure 2.1 shows a representative sparse matrix for a small germanium nanocluster. Also, unlike plane wave methods, the real-space method does not rely on many fast Fourier transforms (FFTs) since calculations are not performed in momentum space. Since FFTs involve global communications, which are hard to parallelize, the real-space method can take better advantage of supercomputer architectures and tend to be faster than plane wave methods for large-scale computations.

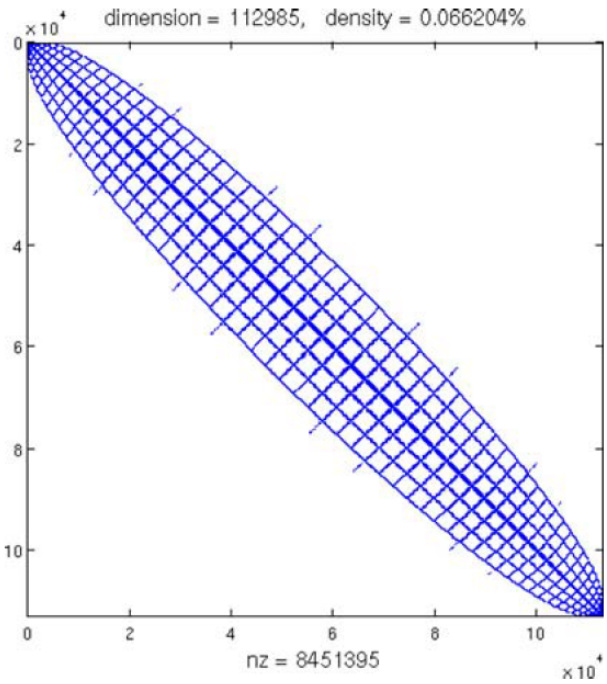


Figure 2.1: Hamiltonian matrix for $Ge_{99}H_{100}$ nanocluster [103].

The kinetic energy term $-\frac{\hbar^2}{2m}\nabla^2$ in the Hamiltonian is calculated on the real-space grid through a finite difference discretization scheme using a higher-order expansion for the Laplacian. This means that the Laplacian is calculated using several neighbors such that:

$$\nabla^2\psi(x,y,z) = \frac{1}{h^2}\sum_{k=-M}^{M}\sum_{l=-M}^{M}\sum_{m=-M}^{M}[C_k\psi(x+kh,y,z) + C_l\psi(x,y+lh,z) + C_m\psi(x,y,z+mh)], \tag{2.7}$$

where M is half of the expansion order and h is the grid spacing. In our tests, we find we achieve a good balance between computational time and accuracy using $M = 6$.

The next term is the ionic potential $V_{ion}(\mathbf{r})$ given by:

$$V_{ion}(\mathbf{r}) = -\sum_{i=1}^{N} \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}|}, \tag{2.8}$$

where i is a label for the N atom system. Calculating this term on the real-space grid runs into some issues. First, the equation contains singularities at atomic sites, which introduces numerical complications. Also, since the potential is steeply diverging, it can cause rapid oscillations of the wave function near atomic sites, which necessitates using a large number of grid points to sample accurately, driving up computational costs. These issues can be dealt with by using pseudopotentials, which will be discussed in the following section.

The standard form of the Hartree potential V_H is given by:

$$V_H(\rho) = \int d^3\mathbf{r}' \rho(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|}. \tag{2.9}$$

Written as is, the Hartree potential is cumbersome to compute since it requires performing an integral for every point $\mathbf{r}$. It is much more practical to solve the Poisson equation:

$$\nabla^2 V_H(\mathbf{r}) = -4\pi\rho(\mathbf{r}). \tag{2.10}$$

We solve this by using the conjugated gradient method and applying various boundary conditions for the potential [15] (we will explore aspects of setting the boundary conditions in the following chapter about nanostructures).

The final term in the Hamiltonian $V_{xc}[\rho(\mathbf{r})]$, the exchange-correlation potential, is also the most controversial term in density functional theory. In principle, the correct exchange-correlation functional exists, which allows for the exact solution of the ground state properties of the system. However, the functional is currently not precisely known, and even if it were, it might take a form so complicated it would be computationally intractable. As previously mentioned, in the Hartree-Fock method calculated energies are always too high because the theory does not fully account for the electron correlation energies. However, the way Hartree-Fock is constructed provides a linear—albeit computationally inelegant—path to the exact solution since one only needs to recover these correlation energies, which can be done by including more basis functions into the wave function. On the other hand, density functional theory provides little guidance on how the functional is to be constructed or even interpreted. Regardless of these issues, density functional theory has proven to be remarkably accurate and remains a widely used research technique owing to its speed and elegance in setting up the electronic structure problem.

The general form of $V_{xc}[\rho(\mathbf{r})]$ is given by:

$$V_{xc}[\rho(\mathbf{r})] = \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})}, \tag{2.11}$$

where E_{xc} can take on various forms based on the approximation method used. One of the most common and easiest-to-implement methods is the local density approximation (LDA) in the Ceperley-Alder form [16]. Here, exchange-correlation effects are similar to those of a uniform electron gas. Another

popular functional is the generalized gradient approximation (GGA) in the Perdew-Burke-Ernzerhof form [99]. In this slightly more complicated functional, exchange-correlation energies rely on not just the local density but also the gradient of the density. Both of these exchange-correlation functionals, along with others, are implemented in PARSEC.

Now that we have described how PARSEC handles each term in the Hamiltonian, we will see how the solution to the Kohn-Sham equation is actually computed. Figure 2.2 shows a schematic for the self-consistent algorithm used in PARSEC. The first step in the algorithm is to obtain an initial guess for the Kohn-Sham potential V_{eff}. A good first guess can be obtained by superposing the atomic charge densities. We then solve the Kohn-Sham eigenvalue problem through standard diagonalization techniques to obtain a new charge density ρ. We use the new ρ to create updated values for V_H and V_{xc}. We combine the updated terms to obtain a new Kohn-Sham potential V_{eff}. Generally, a fraction of the old potential is mixed in with the new for smoother and faster convergence. The new and old V_{eff} are compared by taking their difference. If the difference falls below a desired tolerance, self-consistency has been reached, and the calculated eigenvalue spectrum represents the solution to the Kohn-Sham equation. If the desired tolerance has not been met, the new V_{eff} is tossed back into the loop to obtain a new charge density ρ, which can be used to update V_{eff}, and so on, until convergence is reached.

The most time-consuming part of the algorithm is the diagonalization step needed to solve the eigenvalue problem. Repeated diagonalizations in the

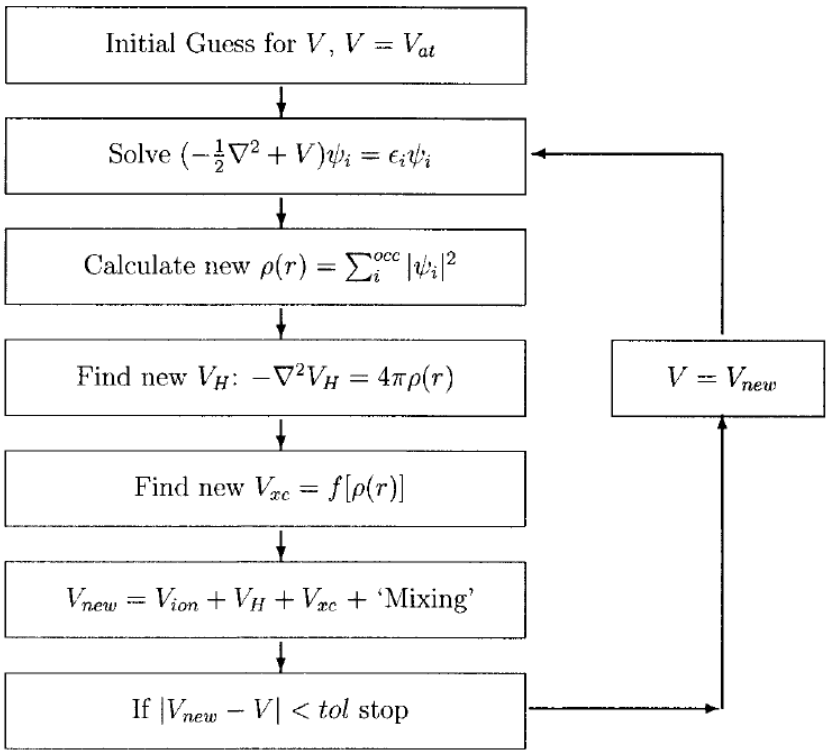


Figure 2.2: Schematic of the self-consistent field loop [103].

loop quickly become costly. As such, we use a Chebyshev filtering method to accelerate the algorithm [132, 133]. With the method, the full diagonalization is only performed once to obtain a good initial approximation for the eigenvalues and eigenvectors. For subsequent loops, diagonalization is replaced by a much less costly subspace filtering operation to update the eigenvalues and eigenvectors. The time savings come from the fact that intermediate terms do not need to be computed to high accuracy, so a full diagonalization is unnecessary. The filter works by honing in on the most crucial terms (the occupied states) and improving them so they progressively resemble the final solution. Self-consistency is generally reached with a similar number of steps as in routines with repeated diagonalizations. Performance tests with the Chebyshev filtering method have shown it can be an order of magnitude faster than these standard diagonalization algorithms [132, 133].

2.4 Pseudopotential theory

Pseudopotential theory was developed to pare down the complexity of atomic potentials. The main idea is to separate the many-electron atom into core and valence states. Core electrons are characterized as being tightly bound to the nucleus. As such, they do not contribute much to bonding behavior. Rather, it is the outermost valence electrons that interact with other atoms and therefore dictate the "interesting" properties of the atom. Pseudopotential theory replaces the inert core electrons with a core potential, leaving only the valence electrons to be computed. This can dramatically reduce the size

of a computational problem. Consider a small nanocluster composed of 100 germanium atoms. An ordinary germanium atom contains 32 electrons, so we would need to include 3,200 electrons to fully describe the nanocluster. With pseudopotential theory, we consider only the valence electrons of which germanium has four. The size of the system shrinks to 400 electrons, which is almost an order of magnitude smaller. The time-savings gained from using pseudopotentials should not be understated: pseudopotential theory is almost single-handedly responsible for the growth of the computational materials field because it made the study of larger systems feasible. It remains a widespread and powerful technique used in all varieties of electronic structure codes.

While pseudopotentials are intended to replicate the behavior of the all-electron potentials they are replacing, in practice pseudopotentials are often constructed to match experimental measurements of key properties. As such, pseudopotential theory is sometimes referred to as being "semi-empirical" in nature. Various methods have been developed for constructing reliable pseudopotentials. Hamann *et al.* outlines four criteria for a desirable pseudopotential for a given atomic configuration [49]. These criteria are as follows:

1. The valence electron eigenvalues for the all-electron potential and the pseudopotential should agree; that is, $\epsilon^{ae}_{val} = \epsilon^{ps}_{val}$.

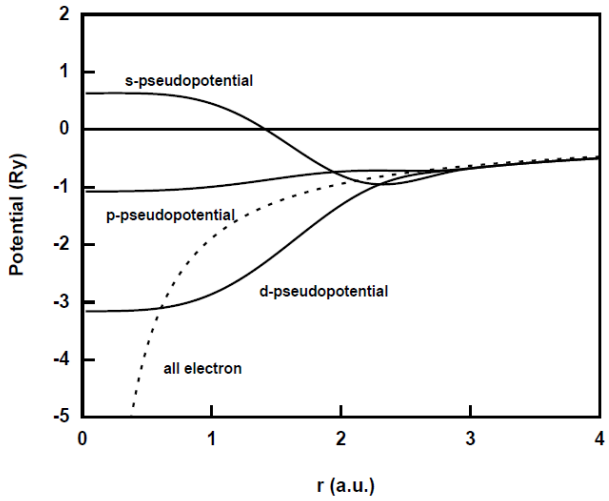
2. The all-electron and pseudo wave functions should agree beyond a chosen cut-off radius; that is, $\psi^{ae} = \psi^{ps}$ for $r > r_c$.

3. The integrals for the all-electron and pseudo charge densities should agree beyond the chosen cut-off radius. This criterion ensures that pseudopotentials are "norm-conserving."

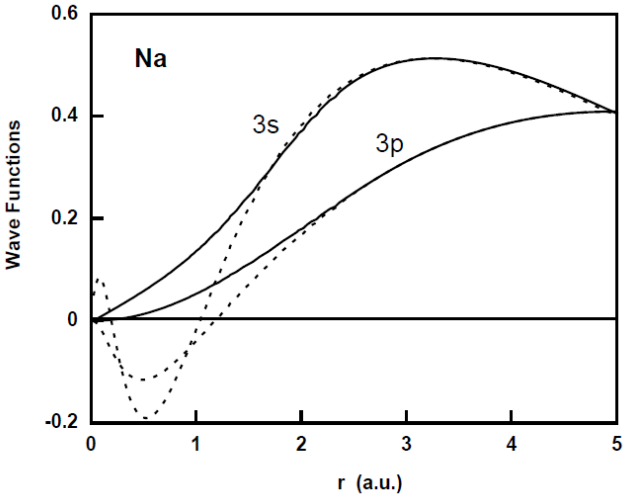
4. The logarithmic derivatives of the all-electron and pseudo wave functions as well as their first energy derivatives should agree beyond the cut-off radius.

The third and fourth criteria are important for transferability. If a pseudopotential has good transferability, it should give accurate results in all chemical environments. This eliminates the need to regenerate and retest pseudopotentials for each and every chemical system. Figure 2.3 shows an example of a pseudopotential and pseudo wave function for a sodium atom. The pseudopotentials and pseudo wave functions match the all-electron cases beyond a user-specified cut-off radius. Pseudopotentials with smaller cut-off radii are referred to as "hard" and require more grid points to sample accurately and reach convergence. Conversely, pseudopotentials with larger cut-off radii—"soft" pseudopotentials—are easier to converge but have worse transferability. In the current research we construct norm-conserving pseudopotentials by the Troullier-Martins scheme [119], which generally gives smooth pseudopotentials with a good balance of transferability and convergence.

Pseudopotentials have a flexible construction and can be tuned for various situations. For example, if there is not a clear separation between the core



(a)



(b)

Figure 2.3: (a) Pseudopotential and (b) pseudo wave function for a sodium atom. The dashed lines represent the all-electron potential and wave functions [24].

and valence states, we can add in "non-linear core corrections" [78]. Pseudopotentials can also be used to account for relativistic effects. A final benefit of using pseudopotentials is computational elegance; compared to the standard ionic potential given by Eqn. 2.8, the pseudopotential does not contain any Coulomb singularities.

One complication involved with using pseudopotentials is that the ionic potential becomes a non-local operator. Orbitals with different angular momentum states (e.g. s, p, d orbitals, and so on) have different scattering behavior and thus require separate pseudopotentials. Non-local projectors are used to put the ionic potential together as a whole. Non-local operators present a computational issue because their matrices tend not to be sparse. So far we have made sure to keep all the terms in the Hamiltonian sparse for the sake of computational efficiency. In order to deal with this, we build our pseudopotentials using a projection scheme proposed by Kleinman and Bylander [63]. The projection scheme writes the pseudopotential in a "semi-local" form by selecting an angular momentum state to act as the local component. In computational tests, the Kleinman-Bylander form of the pseudopotential has been shown to scale well with the size of the basis set and remains a popular choice for constructing pseudopotentials.

Chapter 3

Applying the real-space method to nanostructures

3.1 Introduction

The previous chapter focused on the theoretical foundations of the electronic structure problem. In this brief chapter we will see how these foundations can be applied to the study of nanostructures. Nanoscience is a quickly developing field in materials research owing to the potential for novel applications of nanostructures. We will show how PARSEC's real-space formalism is well-suited to carry out calculations on nanostructures. The key factor lies in how the code implements flexible boundary conditions.

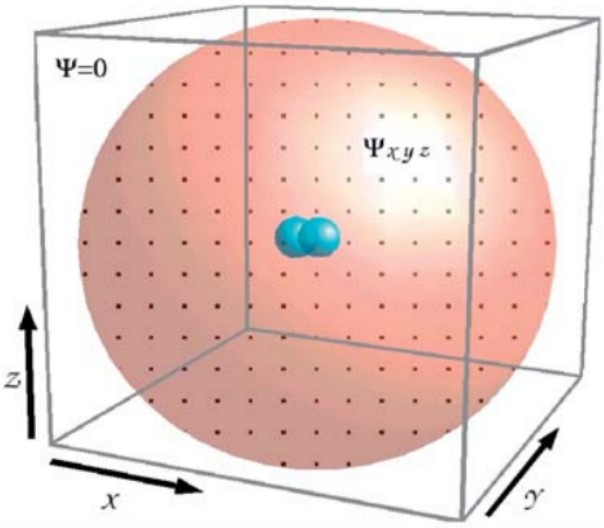
3.2 Periodic boundary conditions

A major benefit to using the real-space method is that it allows for partially periodic boundary conditions. This allows nanostructures to take on a periodicity that naturally suits their geometry. The dimensions in which a system is not periodic represent the confinement directions, where wave functions are required to vanish outside of the domain geometry. For example, nanoclusters and quantum dots are not periodic in any dimension and so are

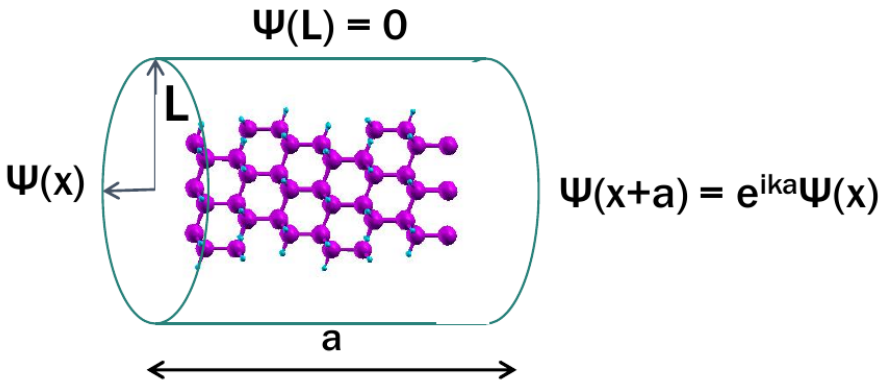
best described with 0D boundary conditions (Fig. 3.1(a)). The real-space grid is set in a spherical domain of a specified radius, outside of which the wave function vanishes. Nanowires and nanoribbons are periodic in the dimension along their axis and so are best described with 1D boundary conditions (Fig. 3.1(b)). The grid is set in a cylindrical domain of a specified radius. Surfaces and sheets are periodic in the lateral planar directions and so are best described with 2D boundary conditions. The domain is a rectangular slab of a specified height. Finally, bulk crystal structures are periodic in all three dimensions, so 3D boundary conditions are used, and the wave function is not confined in any direction.

Most plane wave implementations are restricted to the use of 3D periodic boundary conditions. This can lead to some awkwardness when studying systems that are not periodic in all three dimensions. For example, since a nanowire is only periodic along its axis, the enforced periodicity in the perpendicular directions will result in the creation of a periodic array of nanowire images. The unit cell must be constructed large enough so that a nanowire does not interact with its periodic images, that is, the nanowire does not artificially interact with itself. This may require the use of a very large supercell and involve the calculation of a lot of "wasted space," which is computationally inefficient.

For the sake of simplicity, we only present the formalism for 1D periodic boundary conditions. The discussion can naturally be extended to 2D and 3D periodic boundary conditions [2, 51, 90]. In the 1D case, we consider a system



(a)



(b)

Figure 3.1: (a) System with cluster (0D) boundary conditions. Wave functions vanish outside of a spherical domain of a specified radius [69]. (b) System with wire (1D) boundary conditions. The wave function has a periodicity of a along the axis of the wire and vanishes outside of a cylindrical domain of radius L.

periodic in the x direction with a periodicity of a. The y and z directions are confined by a cylindrical domain of radius L, outside of which the wave function vanishes. From Bloch's theorem, we can write the eigenstates ψ as a plane wave multiplied by some periodic function u:

$$\begin{aligned} \psi_{n,k}(\mathbf{r}) &= e^{ik\cdot\mathbf{r}} u_{n,k}(\mathbf{r}) \\ u_{n,k}(\mathbf{r}+a) &= u_{n,k}(\mathbf{r}) \end{aligned} \tag{3.1}$$

where n is the band index and $\mathbf{k}$ is a wave vector. With the addition of the wave vector k, the Kohn-Sham equation becomes:

$$\begin{aligned} \left(\frac{\hbar^2}{2m}\left(-\nabla^2 + \frac{2}{i}k\frac{\partial}{\partial x}\right) + V_{eff}(\rho)\right) u_{n,k}(\mathbf{r}) &= \epsilon_{n,k} u_{n,k}(\mathbf{r}) \\ \rho(\mathbf{r}) &= \sum_k \sum_{n=1}^{N} |\psi_{n,k}(\mathbf{r})|^2. \end{aligned} \tag{3.2}$$

In order to obtain converged results, we must sample enough k-points in the calculation. This is especially important for accurately capturing the dispersion effects in energy band structures. In our calculations we use a k-point mesh in the form proposed by Monkhorst and Pack [88].

Calculating the ionic potential in periodic systems runs into some issues. The local component of the pseudopotential has a long-range Coulombic behavior, which causes the total energy of the periodic system to diverge. In order to deal with this, we use an Ewald-like technique to sum over the periodic images [35]. We add a compensating Gaussian charge at atomic sites to cancel out the long-range divergence in the local part of the pseudopotential.

To balance out this addition, we must then subtract a corresponding compensating charge to the Hartree potential V_H. The modified V_{ion} and V_H can then be calculated independently, and the Kohn-Sham equation can be solved in the self-consistent loop as usual.

3.3 Charged systems

As with the ionic potential, studying charged systems with periodic boundary conditions can pose problems. The long-range Coulomb force of the charge contributes an artificial interaction via its periodic image and causes the total energy to diverge. For 3D periodic supercell calculations, a popular technique to study charged systems is to introduce a uniformly charged background called a jellium to screen the long-range interaction of the charge. However, the jellium itself can interact with the system to introduce artificial interaction energies. This is particularly true when supercells are used for systems that are not naturally 3D periodic, such as nanowires and surfaces. The jellium is supposed to mainly fill the atoms of the nanostructure to screen its charge. With the supercell approach, the jellium is forced to fill large swaths of vacuum space, which might not be a physically accurate representation of the charged system. Numerous schemes have been devised to correct for the artificial jellium interactions, including the well-known Makov-Payne scheme [80, 70, 37, 114]. The scheme corrects for the total energy by subtracting out the Madelung energy, which represents the spurious electrostatic interaction between the jelium and the charged system. The Madelung

energy takes the form $\frac{\alpha Z_{net}^2}{2\epsilon a}$, where α is the Madelung constant, Z_{net} is the total charge of the system, ϵ is the dielectric constant of the system, and a is the lattice parameter. The Makov-Payne scheme typically requires large supercells to obtain converged results [102]. In addition, for systems with confined periodicities such as nanowires and slabs, the dielectric constant becomes a tensor, which may require additional considerations to determine.

The real-space formalism of PARSEC with its option for partially periodic boundary conditions is naturally equipped to deal with charged systems. The key is in the confinement directions, which allow us to specify an appropriate boundary condition for the potential of a charged system. For a fully confined, 0D periodic system, charged systems pose few issues since naturally, there are no periodic images that cause the energy to diverge. We will discuss the formalism for charged 1D systems. Similar techniques can also be used to deal with charged 2D systems [19].

As mentioned in the previous section, for periodic systems we solve for the compensated Hartree potential $V_H - V_{com}$. We set a boundary condition for this potential based on the multipole expansion for a cylinder in order to include the effects of charge. The monopole term for the potential on the surface looks like:

$$V_{0,0} = \frac{2e}{a} Z_{net} \left[\ln 2 - \gamma - \ln \frac{L}{a} - (\lim_{2\pi p \to 0} \ln(2\pi p)) \right] \tag{3.3}$$

where e is the charge of an electron, Z_{net} is the net charge on the system,

γ is the Euler constant, and p is an integer. Note that if the net charge on the system is 0, the monopole does not contribute to the potential, and $V_H - V_{com}$ takes its ordinary form. If we get rid of the logarithmic divergent term $\lim_{2\pi p \to 0} \ln(2\pi p)$, the potential looks like it corresponds to that of a nonperiodic system. Thus, by eliminating the divergent term, we effectively remove the spurious electrostatic interaction caused by periodicity.

Figure 3.2 demonstrates the 1D charged system formalism on a small phosphorus-doped silicon nanocluster. Although the nanocluster is not naturally periodic, 1D periodicity was artificially imposed so its results can be compared to that of the confined 0D case. The figure shows that our method can be used to correctly calculate ionization energies. Another perk from using the real-space formalism for charged systems is that the reference energy is well-defined, so the eigenvalue spectrum for the charged 1D system is properly aligned with respect to vacuum. No further correction from potential alignment methods is neccessary [70, 37].

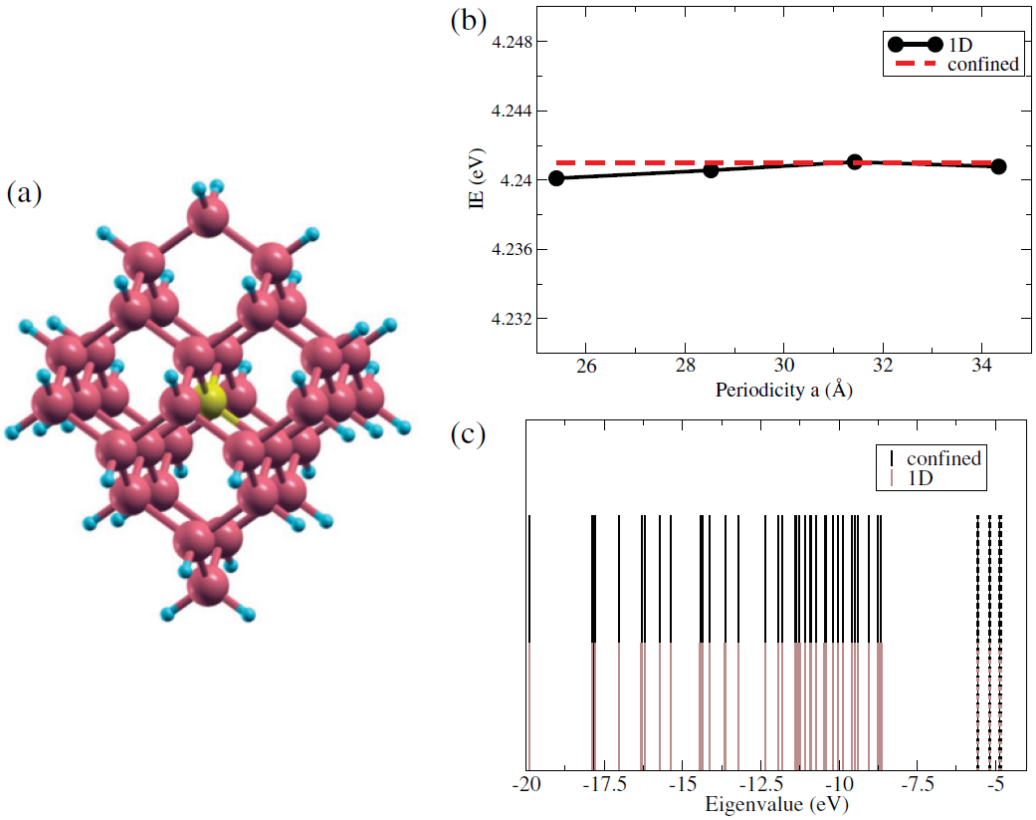


Figure 3.2: (a) Ball-and-stick model of a $Si_{34}H_{36}P$ nanocrystal. Red, yellow, and cyan atoms represent silicon, phosphorus, and hydrogen, respectively. (b) Ionization energy (IE) of the Si nanocrystal as a function of periodicity for an arbitrarily imposed 1D periodic cell. The dashed line is the same structure calculated with 0D (confined) boundary conditions. (c) The eigenvalue spectrum for the +1 ionized Si nanocrystal calculated with different boundary conditions. Solid lines represent occupied states, and dashed lines represent unoccupied states. Red lines were calculated with 1D periodic boundary conditions, and black lines were calculated with 0D periodic boundary conditions [22].

Chapter 4

Mechanical and electronic properties of strained Ge nanowires

4.1 Work summary

Theoretical calculations with real-space pseudopotentials constructed within density-functional theory are employed to calculate mechanical and electronic properties for [100], [110], and [111] germanium nanowires up to 2.7 nm in diameter. Uniaxial strain is applied to wires within the range of -5 to 5%. The strain energy is used to calculate the Young's modulus for each wire, whose values are found to increase with diameter up to approximately the theoretical bulk values. Electronic band structures are calculated for each wire with respect to strain, and from these structures band gaps are obtained. The size and the nature (direct or indirect) of the band gaps are influenced by the growth direction, wire size, and strain amount. Carrier effective masses are calculated from the band structures and jump sharply under certain conditions of strain owing to band crossing, which can correspond to sudden drops in carrier mobilities in applications.

4.2 Introduction

Semiconductor nanowires have sparked recent research interest owing to their unique properties that make them suitable for use in device applications such as photovoltaics and photodetectors [96, 4]. While Si has been the most widely used and studied material in the semiconductor industry, Ge has some advantageous properties over Si such as a smaller band gap and carrier effective masses [112], which can correspond to higher carrier mobilities and better device performance. Studies show that nanowires with higher carrier mobilities are less sensitive to surface roughness scattering [110]. Also, quantum confinement effects may be more pronounced in Ge compared to Si owing to the larger Bohr exciton radius of 24.3 nm in Ge compared to 4.9 nm in Si [79, 30], making Ge potentially more flexible for tuning properties. Some unique applications for Ge include near-IR photodetectors, where Si cannot be used owing to its larger band gap [68],and multijunction photovoltaics like the InGaP/GaAs/Ge solar cell, which has the highest reported efficiency in current solar cell technologies [113].

Compared to bulk materials, which are often brittle except at very high temperatures, nanoscale structures exhibit enhanced strengths owing to higher surface area to volume ratios. Self-purification in nanostructures decreases the concentration of defects, which results in stronger material [31]. Researchers have synthesized Ge nanowires that show mechanical strengths comparable to those of idealized perfect crystals, which is an improvement over the bulk material by orders of magnitude [92]. Under certain conditions, nanowires

can even show plasticity at room temperature, making for a durable, flexible semiconductor useful in many applications [109].

The properties of Ge nanowires depend on their crystal orientation and can change significantly with size and strain effects. Knowledge of how these effects work can be used to tune properties for functionalization [25]. Experimental groups have synthesized nanowires that show a blueshift in the photoluminescence spectrum with decreasing particle size and propose strain to be the cause [4]. A theoretical study showed how strain can be applied to Si/Ge core-shell nanowires to control band offsets in heterojunctions [56]. Strain has been reported to significantly enhance carrier mobilities in Si field-effect transistors [54, 74]. Strain is also an unavoidable, naturally occurring state in many applications, such as in thermal effects. Oxides formed on the surface of nanowires have been shown to add a compressive strain to the nanowire core, affecting the properties of the wire [122]. While numerous *ab initio* computational studies have been performed on Si nanowires, few have been done for Ge [5, 82, 3], only some of which have examined the effects of strain [128, 77, 67]. Here we examine previously unstudied growth directions and further analyze the effects of axial strain on the mechanical and electronic properties of Ge nanowires.

4.3 Computational methods

Electronic structure calculations were carried out using PARSEC, a pseudopotential code for density-functional theory (DFT) calculations in real-

space without the use of an explicit basis [26, 24, 69, 51]. The pseudopotential used for Ge is an improved Troullier-Martins pseudopotential with valence configuration $4s^2 4p^2 4d^0$ with partial core corrections included and the p local component selected [119]. This pseudopotential has been used in previous studies on Ge nanowires with good results [5]. Exchange correlation was handled with Ceperley-Alder, a local density approximation (LDA) functional [16]. Structural relaxations were performed using the BFGS method [13, 36, 41, 106] with a force tolerance of 0.004 Ry/a.u. (1 a.u. $= 0.5292$ Å) The computational parameters for the boundary radius, grid spacing, and number of k-points sampled varied for each system and were optimized until the total energy converged to within 0.01 eV/atom.

With the Ge pseudopotential, the lattice constant for bulk Ge was calculated to be 10.65 a.u., which is within 0.1% of the thermally corrected experimental value of 10.66 a.u. [28]. The bulk modulus was calculated by fitting to the Murnaghan equation of state [89] and gave a value of 73.3 GPa, which is within 4% of the thermally corrected experimental value of 75.8 GPa [28]. The consistency of these results with the literature values suggests the pseudopotential will give accurate results in electronic structure calculations.

Using 1-D periodic boundary conditions in the axial direction, Ge nanowires with sizes ranging from 6 to 27 Å in diameter were carved from the bulk structure in three growth directions [100], [110], and [111]. In this context, diameter refers to the "effective diameter" D defined as $D = 2\sqrt{\frac{S}{\pi}}$ where S is the cross-sectional surface area, calculated by outlining the centers

of the outermost atoms of the wire. The axial lattice constant is defined as the periodic cell length for translational symmetry along the wire axis. With a being the lattice parameter for the cubic unit cell in bulk Ge, the axial lattice constants for [100], [110], and [111] wires are $a, \frac{a\sqrt{2}}{2}$, and $a\sqrt{3}$ respectively. Active Ge atoms on the wire surface were passivated with H atoms [57]. Figure 4.1 shows cross-sections for some of these wires with labelled surface facets. The [100] wires contain four equivalent surface facets, the [110] wires have six surface facets of two distinct types, and the [111] wires have six equivalent surface facets. Note that the surfaces for the [100] and [111] wires are equivalent types.

Structural relaxation was used to calculate the equilibrium axial lattice constant for each system, the results of which are shown in Table 4.1. The table shows that the axial lattice constants contract slightly ($\sim 1\%$) with decreasing wire size and approach the bulk value with increasing size. The lattice contraction is least significant in the [110] direction and shows no variation in the range of sizes tested. These results differ from previous nanowire studies that show a lattice expansion with decreasing size [77, 128, 110]. The mechanism behind the lattice expansion is thought to be a compressive stress on the wire surface that causes axial expansion by the Poisson effect [108]. The nanowires in our simulation were found to be consistent with the Poisson effect; that is, axial tensile stress causes the cross-sectional area to contract slightly, and compressive stress causes the area to expand. Our nanowires do not contradict the mechanism that is thought to cause lattice expansion in

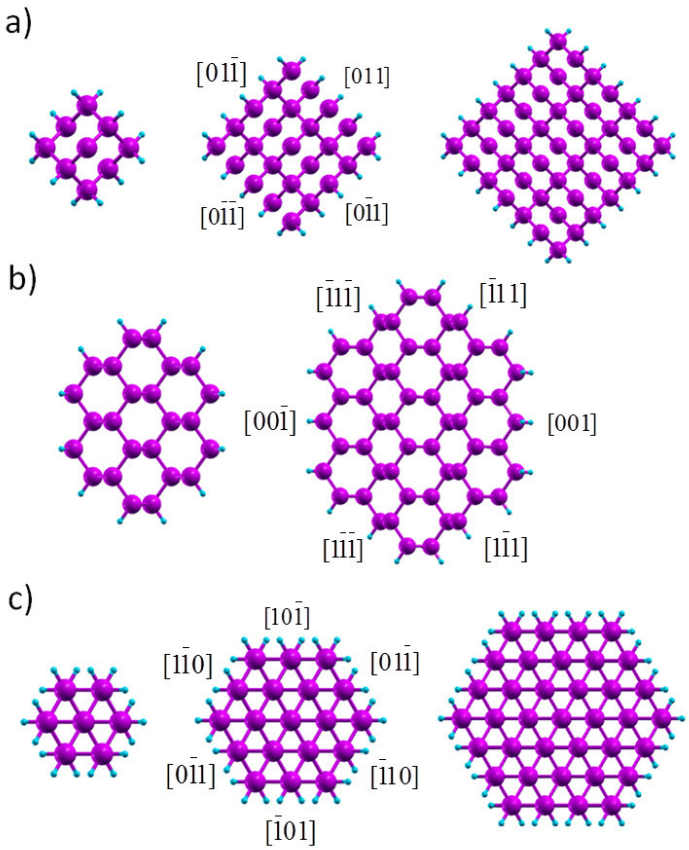


Figure 4.1: Cross-sections with labelled surface facets for a) [100], b) [110], and c) [111] H-passivated Ge nanowires with varying diameters. The surface facets for the [100] and [111] wires are equivalent.

Table 4.1: The number of atoms and axial lattice constants with respect to wire diameter for each growth direction.

Diameter (nm)	# of atoms (Ge; H)	Axial l.c. (Å)
	[100][1]	
0.70	9; 12	5.58
1.2	25; 20	5.59
1.6	49; 28	5.60
2.1	81; 36	5.60
2.5	121; 44	5.62
	[110][2]	
1.30	24; 16	3.98
1.99	54; 24	3.98
2.65	96; 32	3.98
	[111][3]	
0.68	14; 18	9.66
1.09	38; 30	9.70
1.51	74; 42	9.70
1.93	122; 54	9.71
2.35	182; 66	9.72

previous studies, but the lowest energy structures for our wires were calculated to be those where the lattice constant is slightly smaller than that of the bulk. A similar effect has been observed in Si nanocrystals [58].

Strain was simulated by modifying the length of the unit cell in the axial direction in 1% increments and calculating the relaxed structure for each strain step. Wires were strained in the range of -5 to 5%. Experiments have shown that the maximum yield stress for Ge nanowires is around 13% strain, so the range examined in the simulation should have physical meaning [92].

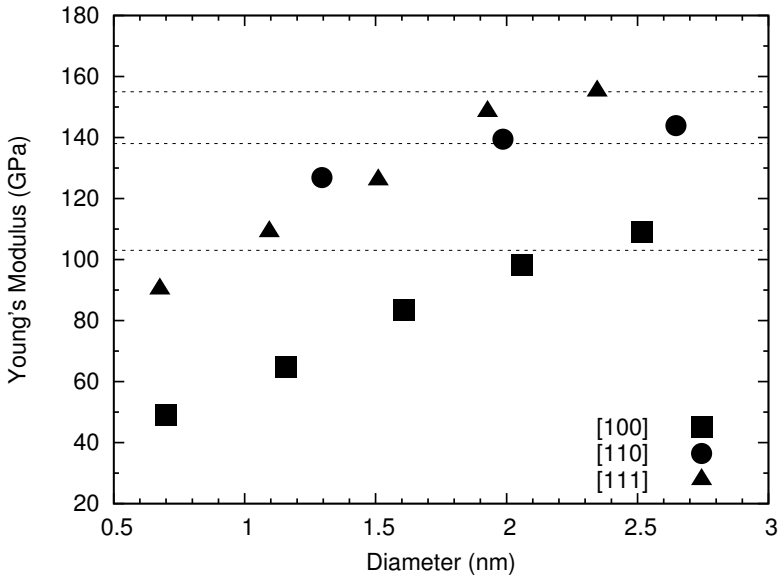


Figure 4.2: Young's modulus vs. diameter for each growth direction. The dashed lines represent the bulk values [9, 125].

4.4 Mechanical and electronic properties

The Young's modulus was calculated for each system from the strain energy curve using second-order polynomial fits of the equation $Y = \frac{1}{V_0}\frac{\partial^2 E}{\partial \epsilon^2}\big|_{\epsilon=0}$ where V_0 is the minimum total energy volume (obtained by multiplying the axial unit cell length with the cross-sectional area S) and ϵ is the strain. Figure 4.2 summarizes these calculations. For all three growth directions, the Young's modulus increases with diameter before appearing to converge at approximately the theoretical bulk values, which are 103, 138, and 155 GPa for the [100], [110], and [111] directions, respectively [9, 125]. Assuming the values are nearly converged, the Young's moduli end up slightly higher than the bulk values. The figure shows that the behavior of the Young's modulus with size

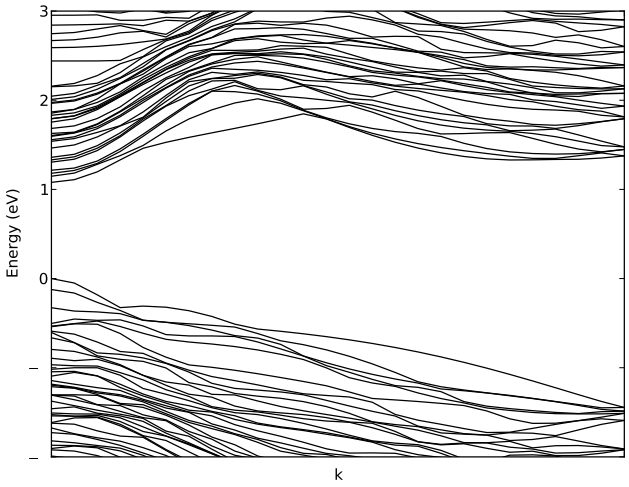


Figure 4.3: Band structure near the band gap for the 1.99 nm wire in the [110] direction. The left edge of k marks Γ, and the right edge marks the BZ boundary. The top of the valence band is set to 0 eV.

is similar for the [100] and [111] wires whereas the [110] wires are noticeably less sensitive. This result mirrors that of the lattice constant contraction, where the [100] and [111] wires showed a similar tendency to contract with decreasing size while the [110] wires hardly varied (Table 4.1). Surface effects can explain these results, as the [100] and [111] wires share equivalent surface facets whereas the [110] wires contain different ones (Figure 4.1). Therefore, if surface effects dominate in the nanoscale, it would be expected that the Young's modulus for [100] and [111] wires behave similarly to each other but different from that of [110] wires.

Band structures were calculated for some of the smaller nanowires in each growth direction. A sample structure is shown in Figure 4.3. Band gaps

were obtained from these structures and are shown in Figure 4.4 and Table 4.2. The band gap for bulk Ge was calculated to be 0.41 eV compared to the literature value of 0.74 eV [112]. While it is well known that DFT within LDA underestimates the absolute magnitudes of band gaps, the general trends for band gaps and carrier effective masses can be reliably reproduced. Studies have shown that band gaps calculated using DFT with LDA scale similarly to more accurate and computationally expensive methods such as GW [124, 131, 91]. The most interesting findings in the carrier effective mass trends, the sudden jumps in value under certain amounts of strain, can be illustrated without absolute accuracy in the magnitudes of the effective masses.

Table 4.2: Band gaps with respect to size for the unstrained wires. The literature value of the gap for bulk Ge is indirect at L with a value of 0.74 eV at 0 K [112].

Diameter (nm)	Band gap (eV)	Direct or indirect? (D or I)
	[100]	
0.70	4.07	D
1.2	2.64	D
1.6	1.92	I
	[110]	
1.30	1.46	D
1.99	1.08	D
	[111]	
0.68	3.38	I
1.09	2.24	D
1.51	1.66	I

The results are similar to those previously reported by other groups [82, 5], which showed direct band gaps for [110] wires and a direct-to-indirect

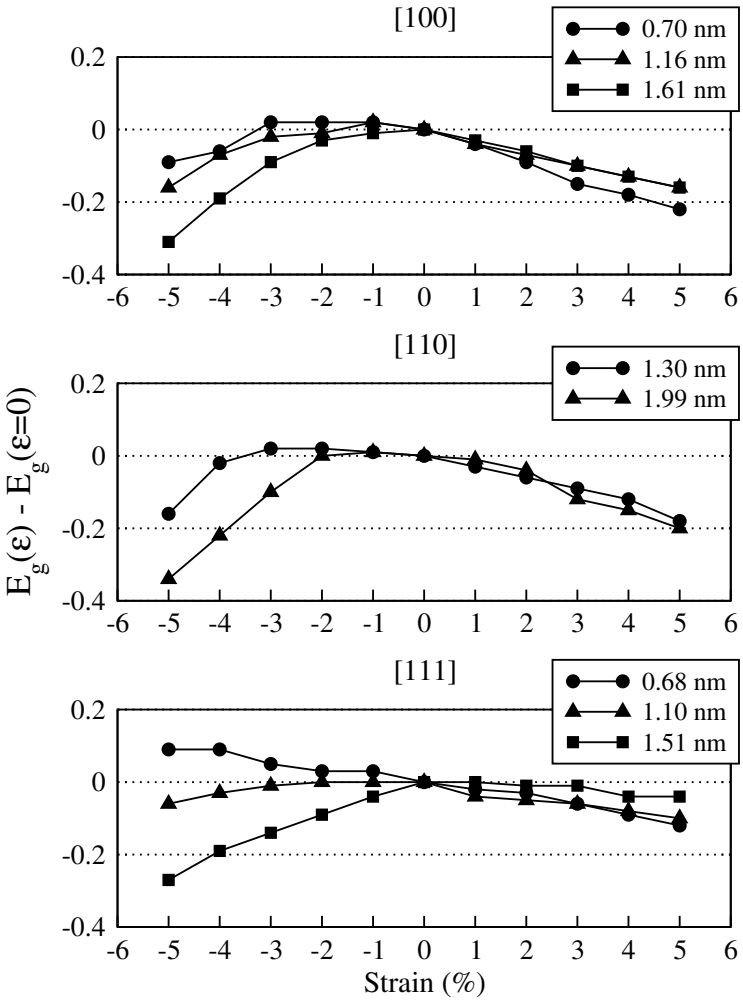
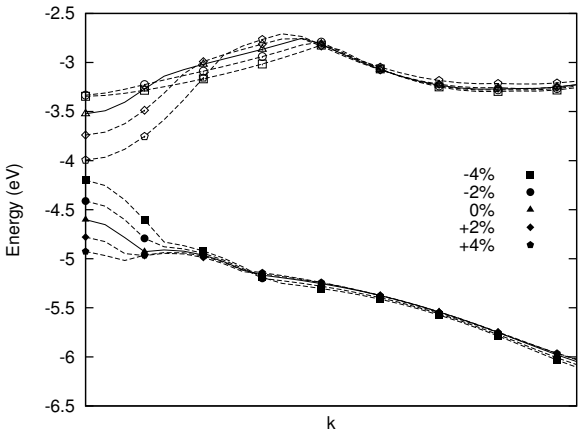


Figure 4.4: Band gap variations with strain, where 0 on the y-axis marks the value for the unstrained cases.

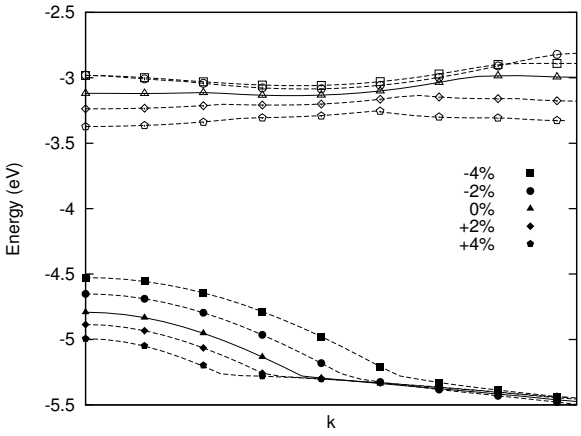
transition for [100] wires above 1.2 nm in diameter. However, for [111], the 1.09 nm diameter wire shows a direct band gap whereas in Medaboina, the band gap is indirect for all sizes. Since the conduction band for the [111] wires has a relatively flat dispersion, the location of the conduction band minimum (CBM) is extremely sensitive to the wire morphology, and whether the band gap can be classified as direct or indirect is often equivocal. Another group has reported direct band gaps for small diameter Ge nanowires in the [111] direction, which contradicts the results reported by Medaboina [100]. But more importantly, the magnitudes of the band gaps are consistent with previous reports and our findings.

The band gaps for Ge nanowires are higher than the bulk value owing to the well-documented quantum confinement effect. The band gaps decrease with increasing wire size and are expected to continue decreasing for larger wires to approximately the bulk value. For comparable wire diameters, band gaps tend to be smallest for [110] wires followed by [111] and then [100] wires. Regarding the effects of strain, all growth directions follow a similar trend, which is a slight decrease of the gap under tension and a slight increase followed by a much steeper decrease under compression. This variance tends to be more sensitive for larger wires. These results are consistent with previous findings on Si and Ge nanowires [110, 82].

Figure 4.5 shows the behavior of the valence and conduction bands with respect to strain for the 1.99 nm wire in the [110] direction and the 1.51 nm wire in the [111] direction. For the unstrained [110] wire, the band



(a)



(b)

Figure 4.5: The valence and conduction bands with respect to strain for the (a) [110] wire with D=1.99 nm and the (b) [111] wire with D=1.51 nm. The left edge of k marks Γ, and the right edge marks the BZ boundary.

gap is direct. Strain affects the shape of the bands more significantly near Γ with relatively little change occuring near the Brillouin zone (BZ) edge. Note that under tension, the energy of the valence band maximum (VBM) at Γ tends to decrease. At 5% strain (not pictured explicitly in the figure), band crossing occurs where the energy of the VBM at Γ drops below the adjacent peak at a slightly higher k-value to which the band gap becomes indirect. Similarly, under compression the CBM at Γ increases and is expected to continue increasing with strain until band crossing occurs with the minimum near the BZ edge.

For the [111] wire, the valence band looks similar to that of the [110] wire, where the VBM occurs at Γ and does not vary much with strain near the BZ edge. The conduction band shows a relatively flat dispersion, which makes the CBM sensitive to wire morphology as mentioned previously. In the unstrained case, the band gap is indirect, but under tensional strain, the energy at Γ can be lowered enough to become a direct band gap. This indirect to direct transition with strain is consistent with previous results [129].

The effective masses for electrons and holes were obtained through second-order polynomial fits around the VBM and CBM using the equation $m^* = \hbar^2(\frac{\partial^2 E}{\partial k^2})^{-1}$. Table 4.3 presents the effective masses for the unstrained cases. The values are consistent with those previously reported for wires in the [100] and [110] direction [77, 6]. The trend shows a decrease in carrier effective masses with increasing wire size.

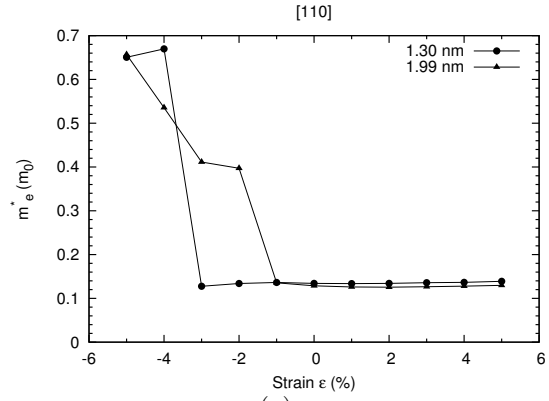
Figure 4.6 shows the behavior of carrier effective masses with respect

Table 4.3: Carrier effective masses with respect to size for the unstrained wires.

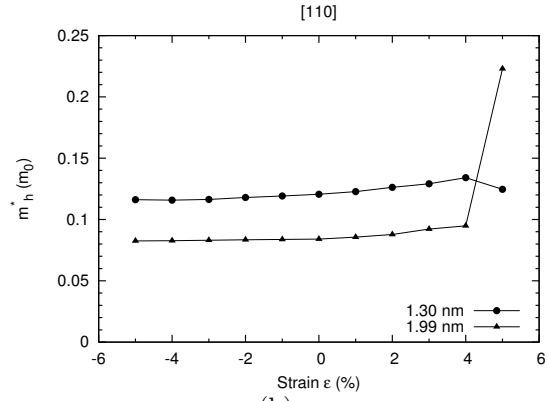
Diameter (nm)	m_e^* (m_0)	m_h^* (m_0)
	[100]	
0.70	0.66	1.89
1.2	0.33	1.06
1.6	0.30	2.18*
	[110]	
1.30	0.13	0.12
1.99	0.13	0.08
	[111]	
0.68	0.86*	0.47
1.09	0.40	0.21
1.51	0.41*	0.14

*Indicates that the VBM or CBM was not at Γ.

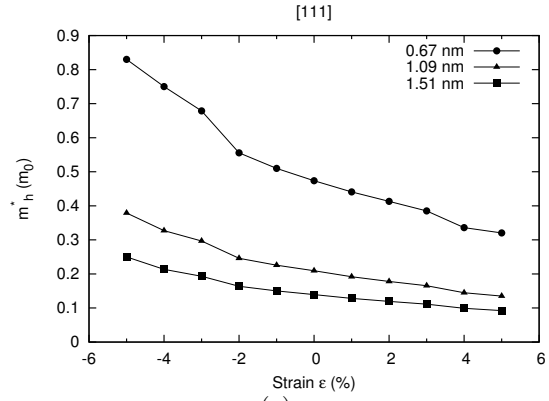
to strain for various systems. In the range of strain tested for the [110] wires, the electron effective mass does not vary significantly with tension. However, under certain amounts of compressional strain, the electron effective mass jumps upward (for the 1.30 nm wire moving from -3 to -4% strain, the electron effective mass jumps by around 425%). This jump corresponds to incipient band crossing that can be mapped to the band structure of Figure 4.5. With compression, the curvature of the CBM at Γ decreases sharply as its energy increases to approach the value of the minimum near the BZ edge. The hole effective mass shows the opposite behavior, where the mass does not vary much with compression, but under significant tensional strain (5% for the 1.99 nm wire), the effective mass jumps. In the band structure, it can be seen that the VBM at Γ starts to drop below the adjacent peak at a slightly higher k-value with 4-5% tensional strain. Since carrier effective masses are inversely



(a)



(b)



(c)

Figure 4.6: Carrier effective masses for (a) [110] electrons, (b) [110] holes, and (c) [111] holes, where m_0 is the free electron rest mass. The jumps in the [110] plots reflect band crossing.

related to their mobilities, these jumps can correspond to sudden decreases in mobilities for certain amounts of strain.

For the [111] wires, the valence band does not show band crossing in the range of strain tested, so no jumps in hole effective masses are observed. The trend shows a slight decrease in hole effective mass with tension and a stronger increase with compression. This can be observed in the band structure where the curvature of the VBM at Γ increases with tension and decreases with compression. The sensitivity of the effective mass with strain seems to decrease with increasing wire size. These results are qualitatively similar to those shown for [111] Si nanowires [110]. The electron effective masses are not included due to the flat dispersion of the conduction band. The effective masses for [100] wires are also omitted as the band structures were plagued with band crossing and did not show any meaningful trends.

4.5 Conclusions

For the three growth directions tested, we find the Young's moduli approach their bulk values with increasing wire size. The behavior of the Young's modulus with size is similar for the [100] and [111] wires, possibly because the surface facets are equivalent in those growth directions. Regarding electronic properties, in the unstrained cases [100] wires show a direct-to-indirect band gap transition with increasing size, and [110] wires show direct gaps for all sizes tested. Tensional strain causes band gaps to decrease slightly whereas compressional strain causes band gaps to first increase slightly then

drop off more steeply. For [110] wires, carrier effective masses jump sharply for certain values of strain, which can correspond to sudden drops in carrier mobilities. For [111] wires, hole effective masses do not show sudden jumps in the range of strain tested. Knowledge of how mechanical and electronic properties change with size and strain can be useful in designing functionalized nanostructures for applications.

Chapter 5

Dopant binding energies in P-doped Ge nanowires

5.1 Work summary

We apply a real-space pseudopotential formalism for charged one-dimensional periodic systems to examine the binding energies of P dopants in Ge[110] nanowires with varying periodicities and diameters. Binding energies calculated by density functional quasiparticle energies of the neutral dopant are severely underestimated whereas those calculated by quasiparticle energies of the ionized defect are overestimated. We find the best method for determining binding energies is to adopt a composite approach that evaluates the total energy difference between charged and neutral systems for the ionization energy of the P dopant, but uses the quasiparticle energy for the electron affinity of the pure Ge nanowire. Our formalism offers a simple density functional method for calculating dopant binding energies of small nanowire systems without the use of computationally intensive many-body perturbation theory calculations.

5.2 Introduction

The dopant binding energy (alternatively referred to in the literature as the dopant ionization or activation energy) is an important property for understanding and evaluating semiconductor performance as it relates to the doping efficiency. For nanostructures, experimental and theoretical studies have shown that the binding energy increases with decreasing size owing to dielectric mismatch and quantum confinement, potentially moving a shallow defect into a deeper state, which can have drastic effects on device functionality [7, 127, 32, 75, 101]. However, significant discrepancies between theoretical models exist, particularly in the magnitude of the binding energy [101, 93], and a more accurate modeling of the defect states will lead to a better understanding of the properties that drives novel technological devices such as functionally doped semiconductor nanowires. Some applications for doped nanowires include biological and chemical sensors, thermoelectrics, and photovoltaics [29, 11, 116].

For a donor in a nanowire, the binding energy E_b is defined as the energy needed to ionize a neutral dopant and move its electron to the conduction band edge far away, evaluated by $E_b = I_d - A_p$, where I_d and A_p refer to the ionization energy and electron affinity of the doped and pure (or intrinsic) system respectively. These two quantities can be calculated from total energy differences between the appropriate charged and neutral systems (ΔSCF). For periodic systems like nanowires, the energy calculation of a charged system is problematic owing to the long-range Coulomb force, which

adds an artificial interaction via its repeated image and causes its total energy to diverge. Typical supercell calculations of charged systems introduce a compensating charge background (or jellium) to screen this long-range interaction. However, the charge background itself interacts with the system and introduces its own artificialities. Numerous schemes have been devised to correct for this interaction to model accurately the energy and potential of a charged material, including the well-known Makov-Payne scheme [80, 70, 37, 114]. These schemes may require very large supercells to obtain converged results, which can be computationally intensive [102].

We will apply an alternative, computationally competitive approach to calculate dopant binding energies in P-doped Ge[110] nanowires. Compared to Si, Ge features a lower band gap and a shallower defect state for P in the bulk [112]. Quantum confinement effects are expected to differ in Ge compared to Si owing to the larger Bohr exciton radius of 24.3 nm in Ge compared to 4.9 nm in Si [79, 30]. Thus, the properties of doped Ge may be more flexible and easily tuned for various electronic applications. Our theoretical approach is based on real-space pseudopotentials constructed within density functional theory (DFT). The charged nanowire is confined within a one-dimensional periodic supercell along the wire axis. Unlike a three-dimensional periodic supercell interwire interactions are eliminated, and total energy corrections related to such interactions can be avoided. In addition it is no longer necessary to include a large amount of vacuum space to converge the total energy of a charged nanowire ($\sim$5 Å is needed), which alleviates the computational

demand. The interaction between image cells along the wire axis can be addressed without introducing a compensating charge background that interacts with the system itself by defining an appropriate electrostatic boundary condition for the Kohn-Sham equation. This approach has been used to examine the capacitance of metallic and semiconductor nanowires [18]. Details of the formalism are outlined in a previous paper [22] .

The dopant binding energy (E_b) can also be obtained by calculating ionization energies (I_d) and electron affinities (A_p) as quasiparticle energies based on the highest occupied (HOMO) and lowest unoccupied (LUMO) Kohn-Sham eigenvalues. This "quasiparticle approximation" is known to have issues within DFT, in part because Kohn-Sham eigenvalues have an ambiguous physical interpretation owing to the delocalization of electrons, and also because DFT underestimates the HOMO-LUMO band gap. Screening charges around the ionized dopant are not accounted for by the HOMO eigenvalue. Considering this shortcoming, Niquet *et al.* [93] proposed that E_b with DFT using the quasiparticle approach can be better approximated by $E_b = A_d(+) - A_p(0)$. Here, $A_d(+)$ is the electron affinity of the ionized doped system obtained from the LUMO eigenvalue after the doped system is ionized, while $A_p(0)$ is still obtained from the LUMO eigenvalue of the corresponding neutral system of pure Ge. That is, I_d is approximated by $A_d(+)$ such that the interaction effects following ionization can be included. In this paper, both the original quasiparticle and Niquet's approach will be evaluated and compared with the ΔSCF approach. For nanowires, while the original quasiparticle approach does not

take into account the interaction between the screening charges and the wire surface, we found that Niquet's approach double counts it. It is best to adopt a composite approach to determine E_b using the total energy difference to calculate I_d but the quasiparticle approach to calculate A_p.

5.3 Computational methods

Electronic structure calculations were carried out using PARSEC, a pseudopotential code for DFT calculations in real-space without the use of an explicit basis set [26, 24, 69, 51]. The pseudopotential used for Ge is based on the Troullier-Martins construction with valence configuration $4s^2 4p^2 4d^0$. Partial core corrections are included and p is adopted as the local component [119]. This pseudopotential has been used in previous studies on Ge nanowires and provides accurate results for the wire's mechanical and electronic properties [72, 5]. Exchange correlation was handled using Ceperley-Alder, a functional based on the local density approximation (LDA) [16]. Structural relaxations were performed using the BFGS method [13, 36, 41, 106] with a force tolerance of $\sim$0.01 Ry/bohr. Only the Γ point was used for k-point sampling as the nanowires examined in the study had relatively large periodicities. Spin polarization was included for systems with unpaired electrons.

Figure 5.1 illustrates one of the Ge nanowires examined in the study. Wires were constructed along the [110] growth direction, as that orientation was shown to be the most energetically stable for small wires [126] and features

some of the more interesting and well-researched electronic properties. Surface atoms were passivated with hydrogen capping atoms [57]. Nanowires were doped by substituting one of the innermost Ge atoms with a phosphorus atom, creating an n-type donor. We constructed and performed calculations on wires of diameter 1.16, 2.28, and 3.42 nm, where the diameter was defined as the smallest cylinder that can enclose the wire.

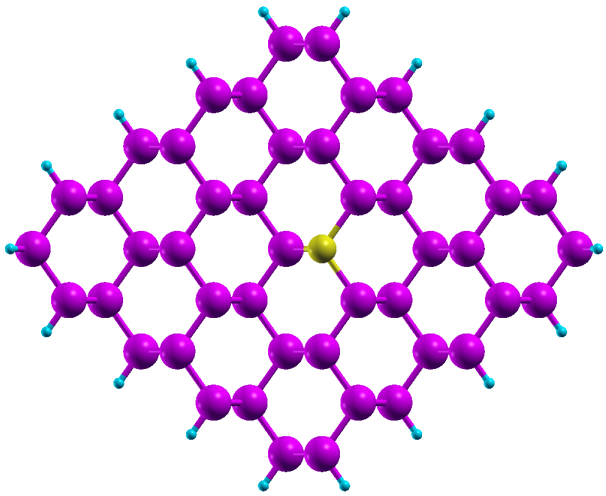


Figure 5.1: Axial view of a H-passivated P-doped Ge[110] nanowire with $D =$ 2.28 nm. The smaller, lighter colored atom near the center is the P dopant. The small atoms at the surface of the wire are hydrogen capping atoms.

5.4 Interaction among P dopants

Since our goal is to study the isolated P dopant, we first examined the dependence of the formation energy E_{form}, ionization energy I_d, and electron affinity A_p on the periodicity of the Ge nanowire. This allows us to assess the

interaction between P dopants from the periodic images along the wire axis. Understanding the behavior of I_d and A_p is particularly important since they determine the binding energy E_b. The nanowire with a diameter $D = 1.16$ nm was chosen for this study. Structural relaxation was not used to ensure trends resulted from changes in periodicity only.

Formation energies were calculated by

$$E_{form}(P) = E(P - Ge) - E(Ge) + \mu(Ge) - \mu(P), \tag{5.1}$$

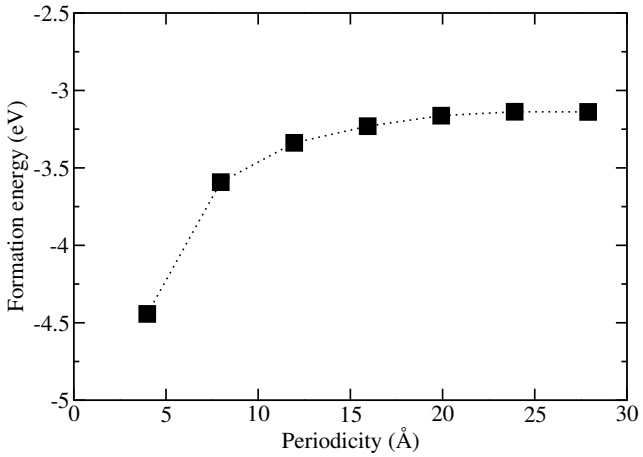
where $E(P-Ge)$ and $E(Ge)$ refer to the total energies of the P-doped and pure Ge wire, and $\mu(Ge)$ and $\mu(P)$ refer to the respective atom's chemical potential. The value for $\mu(Ge)$ was taken to be the energy per atom of bulk Ge, and $\mu(P)$ corresponded to the total energy of an isolated P atom. Similar choices for the chemical potentials were used in previous works on doped nanostructures [50, 23].

In practice the chemical potentials depend on experimental conditions, which may not correspond to the theoretical values adopted here. However, we are only interested in the trend of E_{form}, not in their absolute values. We used a grid spacing of 0.47 a.u., which converges formation energies to within 0.05 eV. The results in Fig. 5.2(a) show that E_{form} increases with periodicity and quickly converges at a periodicity of 23.9 Å. The attractive interaction ($\sim$ 1.3 eV) between P dopants from adjacent image cells quickly diminishes beyond a few Ge bond lengths. Beyond six times the length of the primitive cell, P dopants between image cells are essentially non-interacting, which is

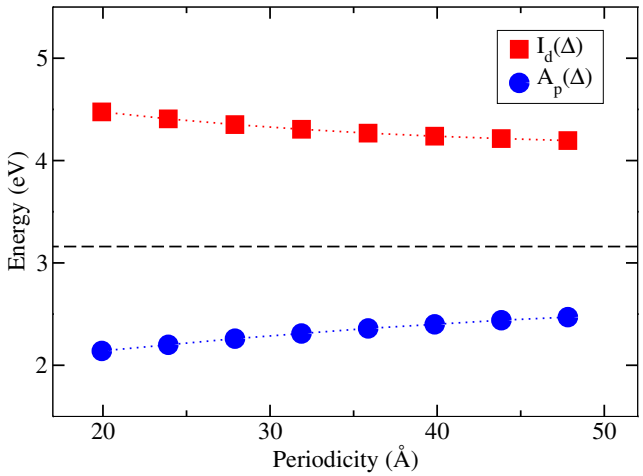
comparable to that in P-doped Si nanowires [23].

Figure 5.2(b) plots the behavior with periodicity for both $I_d(\Delta)$ and $A_p(\Delta)$, with Δ indicating that values were calculated as total energy differences between the appropriate charged and neutral systems. A grid spacing of 0.6 a.u. was used, which converges both quantities to within 0.01 eV. The $I_d(\Delta)$ curve looks similar to that of P-doped Si nanowires [22], which is expected since the ionization energies for Si and Ge nanocrystals were found to be similar [83]. The main characteristic is that $I_d(\Delta)$ increases with decreasing periodicity. Note that since the charged nanowire is enclosed within a cylindrical domain, interwire interactions are eliminated. Furthermore, although the nanowire repeats periodically along its axis, electrostatic intrawire interactions between image cells are also removed by selecting an appropriate reference vacuum level [22]. Therefore, the observed trend in Fig. 5.2(b) corresponds to interactions between unit cells that are non-electrostatic in nature. According to a previous study on P-doped Si[110] nanowires, an induced surface charge forms when the P dopant is ionized. The repulsive interaction between the induced charges in adjacent image cells increases with smaller periodicity leading to a higher $I_d(\Delta)$ [22].

Compared to $I_d(\Delta)$, the $A_p(\Delta)$ decreases concurrently with periodicity. The electron affinity corresponds to the energy gain of a free electron placed in the conduction band minimum (CBM) of a pure Ge nanowire. In contrast to the localized P defect state, the CBM of the pure Ge wire is an extended state. The calculation of $A_p(\Delta)$ should correspond to putting an electron into



(a)



(b)

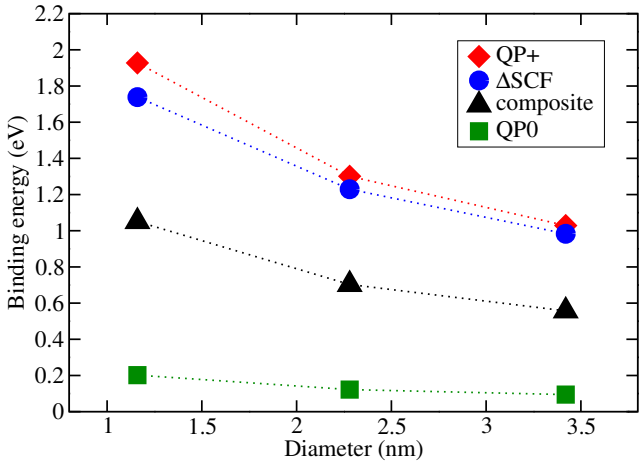
Figure 5.2: (a) The trend of E_{form} of a P dopant in a Ge[110] nanowire ($D = 1.16$ nm) with periodicity. (b) $I_d(\Delta)$ of the same P-doped Ge nanowire in (a) and $A_p(\Delta)$) of the corresponding pure Ge nanowire calculated by total energy differences (ΔSCF). The dashed line corresponds to $|\epsilon_{CBM}|$, which is the magnitude of the eigenvalue of the CBM of the pure Ge nanowire. $A_p(\Delta)$ should converge to $|\epsilon_{CBM}|$ for large periodicities.

a supercell of infinitely large periodicity. Using a smaller periodicity for the calculation thus confines the added electron to a finite region and consequently reduces $A_p(\Delta)$. It is expected that $A_p(\Delta)$ trends to $|\epsilon_{CBM}|$, where ϵ_{CBM} is the eigenvalue of the CBM for the pure Ge nanowire. A DFT study on a large Si nanocrystal suggests that $A_p(\Delta)$ converges to the CBM of bulk Si [132].

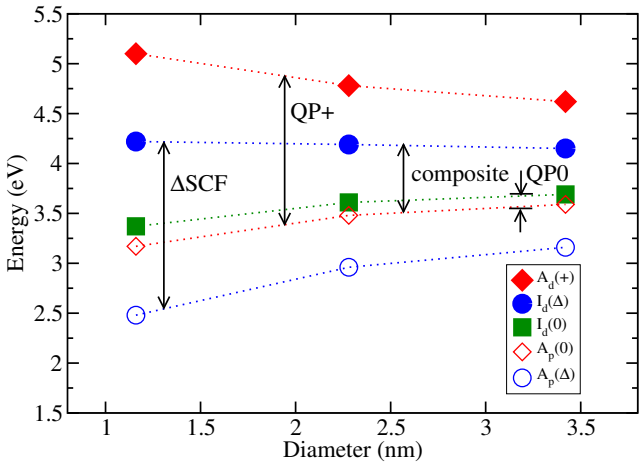
5.5 Quantum confinement effects

The effect of quantum confinement on the binding energy E_b of an isolated P dopant in a Ge[110] nanowire was examined by studying E_b versus wire diameter D. While the interaction between neutral P dopants diminishes beyond ∼24 Å, the interaction between ionized dopants is long ranged. In order to minimize the effect of having a finite periodicity, the wire periodicity is set to be 47.82 Å. At this length, the dispersion of the dopant band is less than 15 meV. Relaxed structures were obtained using a force tolerance of ∼0.01 Ry/bohr.

The binding energy $E_b(\Delta) = I_d(\Delta) - A_p(\Delta)$ can be calculated solely through total energy differences, where both $I_d(\Delta)$ and $A_p(\Delta)$ are calculated from the total energies of the appropriate charged and neutral systems. This approach to calculate binding energies is known as ΔSCF. As discussed in the previous section, $I_d(\Delta)$ includes the quantum mechanical effects following the ionization of P, but these effects are expected to be slightly overestimated due to the finite periodicity of the wire. Figure 5.3(b) shows that $I_d(\Delta)$ is insensitive to D as the curve is quite flat. A similar trend was observed in



(a)



(b)

Figure 5.3: (a) The variation of E_b in a P-doped Ge[110] nanowire with respect to wire diameter D using various methods. (b) The corresponding variations of I_d and A_p with D. The different approaches to calculate E_b are shown in the legend. See the text for details on the methods.

P-doped Si nanocrystals [84, 20] and nanowires [18]. The value for $I_d(\Delta)$ ($\sim$4.15 eV) is similar to that of a P-doped Si nanowire ($\sim$4 eV) [22]. This is to be expected since the ionization energies of Si and Ge nanocrystals are nearly identical [83]. The trend of $A_p(\Delta)$ decreases concurrently with D owing to quantum confinement on the CBM of the pure Ge nanowire. The resultant $E_b(\Delta)$ is plotted in Fig. 5.3(a), which shows an increasing trend with smaller D. The bulk value of E_b for P in Ge is $\sim$13 meV, indicating a a shallow dopant state [81]. For the nanowires examined in the study, the dopants are no longer shallow owing to the effects of dielectric mismatch and quantum confinement as observed in various experimental and theoretical studies on doped semiconductor nanostructures [7, 127, 32, 75, 101].

Since $I_d(\Delta)$ does not vary much with size, the variation of $E_b(\Delta)$ is mainly contributed by $A_p(\Delta)$. Therefore, the accuracy of $A_p(\Delta)$ is crucial as far as the trend of $E_b(\Delta)$ is concerned. Due to the imposed periodicity, $A_p(\Delta)$ was found to be underestimated in the pure Ge nanowire (see Fig. 5.2). As a result, determined values for $E_b(\Delta)$ are too high. It is essential to calculate the ionization energy based on ΔSCF in order to capture the polarization energy following dopant ionization. However, it is more accurate to calculate the electron affinity using a quasiparticle approach such that the effect of imposed periodicity is minimized. The quasiparticle approach calculates A_p using $|\epsilon_{CBM}|$. We denote the resultant electron affinity as $A_p(0)$ to indicate that the eigenvalue was taken from the neutral system. For the range of D examined, $A_p(\Delta)$ and $A_p(0)$ differ by $\sim$ 0.6 eV, but Figure 5.3(b) shows that

the difference narrows with larger D. We suggest a composite approach to calculate the binding energy by $I_d(\Delta) - A_p(0)$. While the effect of quantum confinement is significant in Ge nanowires with small D, the enhancement in the binding energy should not be as large as predicted by $E_b(\Delta)$ but reduced by roughly half as in the composite method.

Both the ΔSCF and composite methods use total energies to compute the binding energy. It is also possible to evaluate E_b using solely a quasiparticle approach. In addition to $A_p(0)$, the ionization energy can be calculated using $|\epsilon_{defect}|$, where ϵ_{defect} is the eigenvalue of the P defect level within the gap of the Ge nanowire. We denote this ionization energy $I_d(0)$. In Figure 5.3(b), $I_d(0)$ decreases with decreasing D as in $A_p(0)$. This behavior can be understood using the hydrogenic model of a shallow dopant in semiconductor [112], which states that the trend of the defect level follows that of the semiconductor's CBM. The corresponding binding energy is $E_b(0) = I_d(0) - A_p(0)$ as illustrated in Fig. 5.3(a). We call this method QP0 to indicate that the ionization energy is calculated using the quasiparticle of the neutral doped system. $E_b(0)$ is severely underestimated compared to the ΔSCF methods owing to DFT's self-interaction error that results in an incorrect description of the energy levels in the conduction band. Similar issues were observed in P-doped Si nanowires using similar methods and functionals [93, 101]. The DFT error is a relatively minor issue in the ΔSCF approach since the total energies of ground states are generally accurate when the system does not involve highly localized electronic states. The underestimate of $E_b(0)$ also originates from the lack of accounting

by $I_d(0)$ for the energy associated with polarization after the dopant is ionized. This should not be a significant concern for $A_p(0)$ since the CBM is not a localized state, so the energy associated with polarization should be relatively small.

Noting the shortcomings of the QP0 approach, Niquet *et al.* suggested an alternative method that uses the ionized doped nanowire instead of the neutral system. They used the LUMO eigenvalue of the ionized system (denoted by $A_d(+)$) as the ionization energy of the neutral nanowire. This way the polarization energy associated with screening following dopant ionization is taken into account. The binding energy thus becomes $E_b(+) = A_d(+) - A_p(0)$. We call this approach QP+ to indicate that the ionization energy is calculated by a quasiparticle approach that uses an ionized dopant. The calculated $A_d(+)$ vs. D in Fig. 5.3(b) shows that the effect of screening is exaggerated by QP+ as $A_d(+)$ values are too high compared to $I_d(\Delta)$.

We found that the $I_d(\Delta)$ values are almost exactly the average of $I_d(0)$ and $A_d(+)$, or

$$A_d(+) \approx I_d(\Delta) + (I_d(\Delta) - I_d(0)) \,. \tag{5.2}$$

If $I_d(\Delta)$ is taken to be accurate, then $I_d(\Delta) - I_d(0)$ corresponds to the error of $I_d(0)$; thus $A_d(+)$, which was supposed to account for this error, ends up double counting it. Using the expression for $A_d(+)$ in Eqn. 5.2, it can be shown that $E_b(+) \approx E_b(\Delta) + (E_b(\Delta) - E_b(0))$. Therefore, Niquet's QP+ method, which was devised to correct for the underestimate of E_b by the QP0 method, turns out to overestimate the binding energies as illustrated in Fig. 5.3(a).

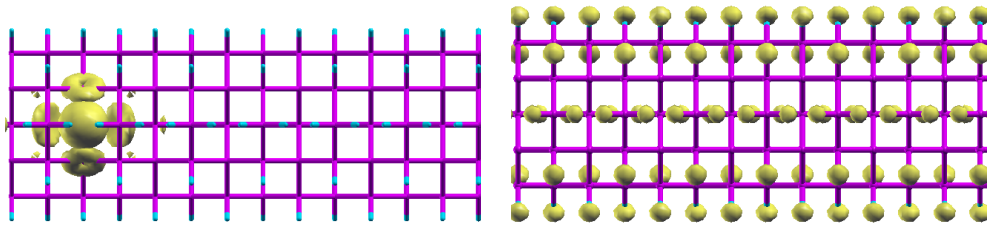
One of the key differences between the QP0, QP+, and ΔSCF methods comes from the treatment of the ionization energy I_d of the doped system. A similar comparison between the magnitudes of I_d using these methods can be found in a study on small Na clusters [14]. In this study, ionization energies were calculated with LDA as well as with many body perturbation theory (MBPT), including self-consistent GW (scGW). In the LDA results, $I_d(0)$ calculated from the HOMO of the neutral species were underestimated, and $I_d(+)$ calculated from the LUMO of the ionized species were overestimated, but $I_d(\Delta)$ calculated as total energy differences gave results closest to experimental values, being slightly overestimated by a few tenths of an eV. The LDA results for $I_d(\Delta)$ are reasonably close to the scGW results, often within 0.1 eV. Another study on small Ge clusters that compares ionization energies calculated by LDA and GW reports similar findings, where the LDA results for $I_d(\Delta)$ are slightly overestimated and those for $I_d(0)$ are grossly underestimated [59]. The study also computes electron affinities and finds that those calculated by LDA with the ΔSCF method are overestimated by a similar amount compared to GW and experimental results.

The composite method $E_b(composite) = I_d(\Delta) - A_p(0)$ thus adopts $I_d(\Delta)$ but continues to treat $A_p(0)$ using the quasiparticle approach, which is appropriate for an extended state (the CBM of the pure Ge nanowire) in order to mitigate the effects of imposing periodicity. The most accurate result for the binding energy should correspond to the composite method, where QP0 underestimates the binding energies while ΔSCF and QP+ overestimate

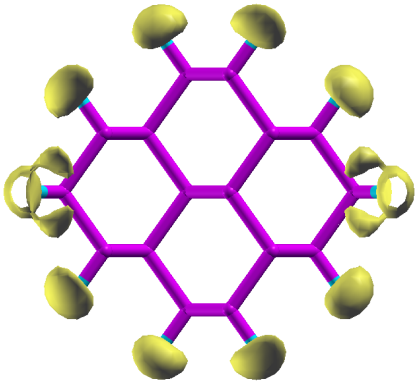
them. The differences between the various methods become less significant for larger wires. Ideally, the various methods should be benchmarked against MBPT since the study of defects involves the energy band gap, which LDA has problems reproducing; however, we could not find a relevant MBPT study for our system.

To attain a deeper understanding of the charged nanowires, we examined the induced charge ρ_{ind} of the ionized P-doped and the negatively charged pure Ge nanowire ($D = 1.16\ nm$). The induced charge density of the ionized P-doped nanowire ($\rho_{ind}(P)$) was obtained by taking the difference between the charge density of the ionized (positively charged) P-doped nanowire and that of the neutral undoped wire. The induced charge density of the negatively charged pure Ge wire ($\rho_{ind}(Ge)$) was obtained by taking the difference between the charge density that corresponds to the first N electrons of the charged pure wire (with $N + 1$ electrons) and that of the neutral pure wire (with N electrons). As spin polarization was included for systems with unpaired electrons, the relative contributions to the charge density of the up and down spin states were weighted by their occupancies. In each case the charged density has the same number of electrons (N) such that ρ_{ind} corresponds to zero net charge. A positive (negative) ρ_{ind} indicates an excess (deficiency) of electrons compared to the neutral system.

Surface contour plots of $\rho_{ind}(P)$ and $\rho_{ind}(Ge)$ are shown in Fig. 5.4. The contour surface corresponds to a chosen positive value of ρ_{ind}. The plots show that $\rho_{ind}(P)$ is localized around the P dopant [Fig. 5.4(a)] whereas $\rho_{ind}(Ge)$ is



(a) (b)

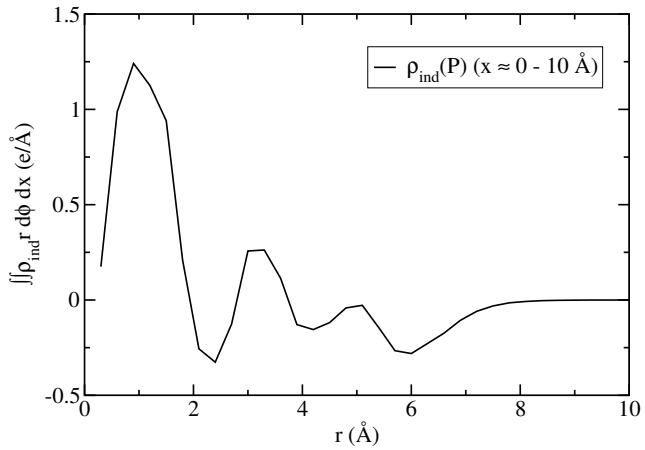


(c)

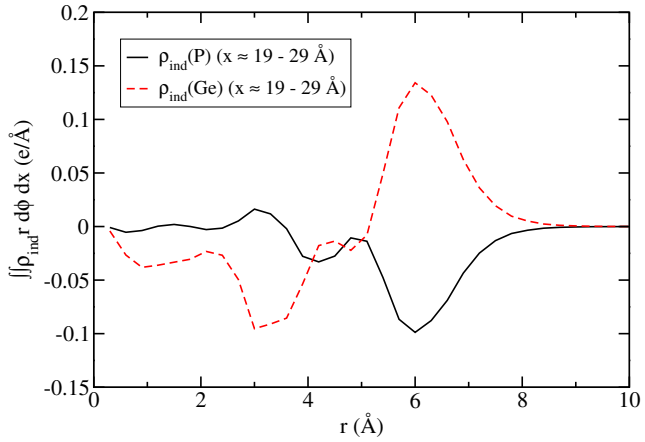
Figure 5.4: Surface contour plots for the induced charge density ρ_{ind} of various Ge nanowires ($D = 1.16$ nm). The contour surface corresponds to a chosen positive value of ρ_{ind}. Only the atomic bonds are drawn: individual atoms are not shown. (a) Side view of $\rho_{ind}(P)$ for the ionized P-doped wire. The $\rho_{ind}(P)$ is localized around the P atom. (b) Side view of $\rho_{ind}(Ge)$ for the negatively charged pure Ge wire. In contrast to (a), $\rho_{ind}(Ge)$ is spread evenly throughout the wire. (c) Axial view of (b). The $\rho_{ind}(Ge)$ is spread around the surface of the wire at $r \approx 6$ Å.

evenly spread throughout the wire axis [Fig. 5.4(b)]. This originates from the fact that the P defect state is localized whereas the CBM of the pure Ge wire is an extended state.

The radial variation of $\rho_{ind}(P)$ and $\rho_{ind}(Ge)$ was calculated in ~ 10 Å segments along the wire axis. Figure 5.5(a) plots a segment that contains the P dopant (located at $x \approx 5$ Å) whereas Fig. 5.5(b) plots a segment of the wire far from the P dopant ($x \approx 19 - 29$ Å). In the ionized P-doped wire, the electrons that screen the positive P ion are mostly drawn from around the surface of the nanowire, reflected in the figure as a deficiency of electrons (negative $\rho_{ind}(P)$) at $r \approx 6$ Å. Since the DFT eigenvalue of the defect level ϵ_{defect} does not capture such polarization effects, it is vital to calculate I_d using the ΔSCF approach. The electron deficiency at the wire surface extends to the segment far from the P dopant although $\rho_{ind}(P)$ decreases in magnitude. The $\rho_{ind}(P)$ on the nanowire surface leads to a repulsive interaction between periodic cells that results in an enhanced I_d as described in Section 5.4. For the charged pure Ge wire, since the electron is added to an extended state, the radial variation of $\rho_{ind}(Ge)$ is independent of the chosen segment along the wire axis (the $x \approx 19 - 29$ Åsegment is plotted). In this case Fig. 5.5(b) shows an excess of electrons around the surface of the nanowire at $r \approx 6$Å. This is also depicted in the surface contour plot of $\rho_{ind}(Ge)$ in Fig. 5.4(c). Note that the $\rho_{ind}(Ge)$ plots correspond to a finite linear charge density along the pure Ge nanowire. In contrast to the ionized P-doped wire, a single electron deposited into the CBM of an infinitely long pure Ge nanowire should not lead to a considerable



(a)



(b)

Figure 5.5: (a) The radial variation of $\rho_{ind}(P)$ for the ionized P-doped Ge wire in a ~ 10 Åsegment along the x-axis that contains the P dopant (located at $x \approx 5$ Å). (b) Same as in (a) but for a segment far from the P dopant ($x \approx 19 - 29$ Å). For the same segment, $\rho_{ind}(Ge)$ for a charged pure Ge nanowire is also plotted for comparison.

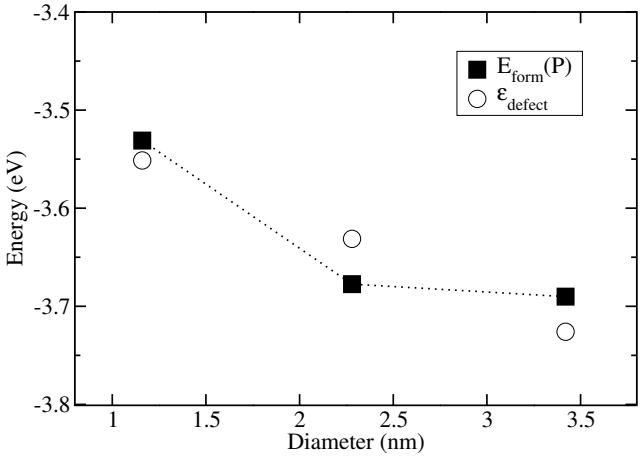


Figure 5.6: The formation energy $E_{form}(P)$ of doping P into a Ge[110] nanowire as a function of wire diameter D. Relaxed structures were used for the calculations. The eigenvalue of the dopant defect level ϵ_{defect}=$-I_d(0)$ is also shown.

ρ_{ind} as indicated. Therefore, the electron affinity should be well approximated by ϵ_{CBM}.

While the effect of quantum confinement on E_b should involve $I_d(\Delta)$, the formation energy $E_{form}(P)$ obtained from relaxed structures can be correlated to $I_d(0)$. Figure 5.6 plots $E_{form}(P)$ as a function of D and shows that the magnitude of $E_{form}(P)$ decreases concurrently with D. Since $|E_{form}(P)|$ represents the energy of replacing a Ge atom with P, the trend implies that the nanowire becomes more difficult to dope as it becomes smaller. A similar "self-purification" effect has also been observed in doped Si nanostructures [50, 31]. The eigenvalue of the defect level ϵ_{defect} (or $-I_d(0)$) is also plotted (it is by coincidence that the ϵ_{defect} are aligned with the $E_{form}(P)$ without introducing a shift). Its trend with D follows that of $E_{form}(P)$ quite well, which provides

evidence that the energetics of a shallow dopant in a nanostructure can be correlated with the Kohn-Sham eigenvalues of the dopant defect level [23].

5.6 Conclusions

We used our real-space DFT formalism for charged one-dimensional periodic systems to calculate and compare the results of the dopant binding energy E_b with respect to wire diameter for a P dopant in a Ge[110] nanowire using various methods. For all sizes of nanowire examined, dopant levels are not shallow as in the bulk. Binding energies calculated using quasiparticle energies (QP0) of the neutral dopant are underestimated owing to the lack of accounting for the polarization energy associated with screening the ion, resulting in too shallow dopant levels. Conversely, the quasiparticle energies of the ionized defect (QP+) overestimate the binding energy because the polarization energy is double counted. Binding energies calculated by total energy differences (ΔSCF) between charged and neutral systems also overestimate the binding energy, and the error originates from the imposed periodicity of the wire. Our proposed composite method calculates the ionization energy $I_d(\Delta)$ of the P dopant by total energy differences, but calculates the electron affinity $A_p(0)$ of the pure Ge nanowire using a quasiparticle approach. Such a composite approach to calculate the binding energies likely gives the most accurate results. The study explores how our formalism for modeling charged 1-D periodic systems without a compensating charge background can be used to provide a computationally efficient DFT method for determining dopant

binding energies in nanowires.

Chapter 6

AFM simulations overview

6.1 Introduction

Developed in 1986, atomic force microscopy (AFM) is a powerful probe-based imaging technique that can be used to visualize and characterize chemical phenomena. Remarkably, this technique has even been able to directly visualize chemical features known mostly to theory, such as bond order [43] and intermolecular bonding [130]. Molecular imaging can also be used as a characterization tool to identify target molecules and structural properties like defects. While probe microscopy is a rapidly developing research field, the interpretation of experimental images is not well-understood and is in some cases controversial. This motivated numerous theoretical studies to use modeling techniques to gain an in-depth understanding the fundamental physics behind the AFM imaging mechanism [42, 46, 86, 95, 43, 52, 53, 48, 111, 45, 47, 62, 34]. These studies have examined effects such as tip functionalization and the types of forces on the tip (Pauli, electrostatic, and van der Waals, for example) and their contribution to the contrast mapping of the image. A comprehensive understanding of the AFM mechanism can be useful for guiding future experiments and driving advancements in imaging technologies.

The basic components of AFM are the sample and the cantilever with a tip (Fig. 6.1). The sample is prepared and mounted on a surface, often

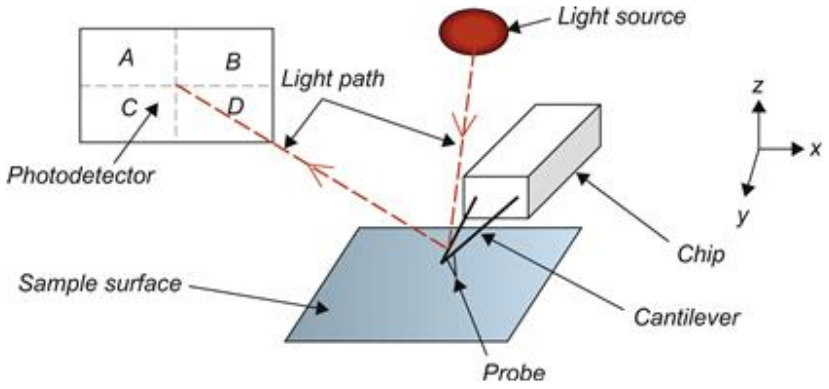


Figure 6.1: Basic AFM schematic [12].

by binding to a substrate, and the tip is scanned in a 2D plane over the sample. In the noncontact AFM mode (nc-AFM) the tip never comes directly into contact with the sample but hovers over it at a height on the order of angstroms. During the scanning process the cantilever deflects based on the tip's response to the sample. This response is detected by a laser and converted into a signal that makes up the AFM image. Experimentalists have discovered that detection sensitivity can be increased by functionalizing the cantilever tip with a probe molecule such as CO. The extraordinary resolution that can be obtained by nc-AFM with CO functionalized tips has been used to visualize the detailed chemical structure of molecules such as pentacene, resolving its individual atoms and chemical bonds [42].

Theoretical AFM simulations can be difficult to carry out because the

morphology of the probe tip and the dynamics during the scanning process are not precisely known. An accurate treatment of the tip is important since its interaction with the sample is responsible for the AFM image. However, constructing the tip in detail can be computationally intensive. Various models, which often rely on the appropriate use of simplifying assumptions, have been developed to make this problem computationally tractable.

A potential application for theoretical imaging simulations is to help characterize experimental results by creating benchmarks for comparison. Currently there is no simple guideline to follow, so the interpretation of experimental images is rather *ad hoc*. This poses a problem when irregularities appear on images. Theoretical simulations can be used to translate distortions or irregularities into known physics. The following work aims is to use *ab initio* electronic structure techniques to develop an accurate and computationally efficient method to create simulated AFM images that can aid the interpretation of experimental images.

6.2 Methods of AFM simulations

In AFM imaging, the contrast mapping is not a direct representation of the force but instead is a measurement of the frequency shift. In the AFM scanning process, we can describe the tip probe as a one-dimensional harmonic oscillator, and the frequency shift Δf measured in AFM can be related to the interaction energy between the tip probe and sample E_{ts} with respect to the tip-sample distance q by the equation:

$$\Delta f = \frac{f_0}{\pi k_0 A^2} \int_{-A}^{A} \frac{\partial^2 E_{ts}(z-q)}{\partial q^2} \sqrt{A^2 - q^2} dq, \tag{6.1}$$

where f_0, A, k_0, and z are the resonant frequency, oscillation amplitude, spring constant, and operating height of the AFM tip respectively. (The derivation of this equation is not trivial and can be found in the references [40].) By applying the small amplitude oscillation approximation to the tip, this equation simplifies to:

$$\Delta f = \frac{f_0}{2k_0} \frac{\partial^2 E_{ts}}{\partial z^2}, \tag{6.2}$$

where the coefficient in front of the derivative is a constant whose value can be neglected since the contrast mapping depends on the relative frequency shift. The main advantage of using the small amplitude approximation is to reduce the number of parameters and data points needed to produce an AFM image. In our testing, we did not see a notable difference between images computed by Eqs. 6.1 and 6.2. In most cases we used the easier-to-implement small amplitude approximation for our simulations.

6.3 Full DFT method

The full DFT method is a conceptually simple "brute force" approach for simulating AFM images. Here, the tip probe is modeled explicitly as a small system, such as the two atom CO or Cu_2 molecule. The probe is then rastered over the sample on a fine 2D grid at varying tip heights z with each raster point

requiring a separate total energy calculation, which represents the tip-sample interaction at that tip location. Then Δf given by Eqs. 6.1 and 6.2 can be obtained by taking the numerical derivative of the total energy calculations. Since this method requires sampling multiple tip heights, with each height requiring a raster grid of total energy calculations, a single AFM image may require thousands of individual calculations to compute. This highlights the importance of having an efficient computational method since savings from one calculation will be multiplied thousands of times over. The computational burden of having to perform such a large number of calculations to create a single AFM image necessitated the development of other AFM methods.

6.4 Frozen density embedding simulation method

Part of the tedium in the full DFT simulation method comes from having to perform a large number of very similar calculations, where only the tip location is changing while the sample stays the same. The unchanging nature of the sample for a weakly interacting tip inspired the development of using frozen density embedding theory (FDE) in AFM simulations as an approximation to full DFT calculations. The main premise of the method is that it approximates the sample as an external field taken from a previous self-consistent calculation. This decreases the size of the system to that of the tip, which results in vast computational savings considering that the sample is often the largest part of the calculation and that the savings are multiplied many times over since computing an AFM image requires repeated DFT calculations.

A detailed description of the formalism is laid out in a previous paper [104], which is summarized here. We divide the total charge density of the tip-sample system $[n^{tot}(\mathbf{r})]$ into two subsystems of the tip and sample: $n^{tot}(\mathbf{r}) = n^t(\mathbf{r}) + n^s(\mathbf{r})$. We assume that the interaction between the tip and sample is small enough that the tip does not affect the structural or electronic properties of the sample. The Kohn-Sham equation for the divided system becomes:

$$\left[\frac{-\nabla^2}{2} + V^t_{eff}(\mathbf{r}) + V_{emb}(\mathbf{r})\right]\psi^t_i(\mathbf{r}) = \epsilon^t_i\psi^t_i(\mathbf{r}), \tag{6.3}$$

where $V_{eff}(\mathbf{r})^t$ is the standard tip potential and $V_{emb}(\mathbf{r})$ is the embedded sample potential. The full expression for the sample potential is given by:

$$\begin{aligned} V_{emb}(\mathbf{r}) = \int \frac{n^s(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + V^s_{nuc}(\mathbf{r}) + \left.\frac{\delta E_{xc}[n]}{\delta n}\right|_{n=n^{tot}} \\ - \left.\frac{\delta E_{xc}[n]}{\delta n}\right|_{n=n^t} + \left.\frac{\delta T_s[n]}{\delta n}\right|_{n=n^{tot}} - \left.\frac{\delta T_s[n]}{\delta n}\right|_{n=n^t}, \end{aligned} \tag{6.4}$$

where V_{nuc}, E_{xc}, and T_s represent the nuclear potential, exchange-correlation energy functional, and kinetic energy functional respectively. The kinetic energy terms can be calculated through the choice of an appropriate kinetic energy functional. In our tests, we find that the functional proposed by Tran and Weslowszki [118], which takes a form similar to that of the Perdew-Burke-Ernzerhof exchange functional [99], gives reasonable results.

Figure 6.2 compares simulated AFM images of pentacene with a CO tip using the full DFT and FDE methods. The images are virtually identical and are also consistent with the experimental image. An important simulation detail to mention is that the full DFT simulation was taken at tip height $z = 3.39$ Å whereas the FDE simulation was taken at $z = 3.07$ Å. In our testing, the FDE method with the Tran-Weslowszki kinetic energy functional tends to overbind, resulting in a consistent underestimation of the intermolecular distance between the tip and sample of about 0.3 Å. This distance can be offset manually in order to give qualitatively accurate images between the full DFT and FDE simulation methods.

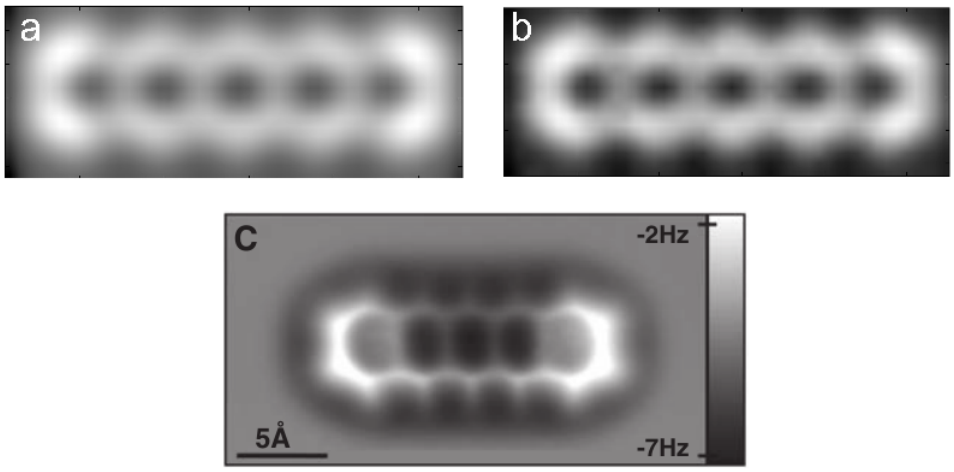


Figure 6.2: Simulated AFM images of pentacene with a CO tip using the (a) full DFT and (b) frozen density embedding (FDE) methods, where tip heights are set to $z = 3.39$ and 3.07 Å respectively [104]. (c) Experimental AFM image [42].

6.5 Virtual tip method

The virtual tip method was developed partly to get around the large numbers of calculations needed to do a full DFT or FDE calculation. Using this method, an AFM image can be calculated from a single self-consistent calculation rather than thousands, making it suitable for use as a quick screening tool. It achieves this by not including an explicit model for the tip in the calculation. This is not to say that the simulation is modeled without a tip. Rather, the tip is approximated as a virtual object in order to measure the forces responsible for the AFM contrast mapping. A comprehensive description of the virtual tip AFM simulation method is detailed in the references [21, 61]. A summary of the main principles and key derivations are provided here.

Assuming the tip does not affect the structural and electronic properties of the sample, the sample can be treated as a perturbation to the tip-sample system and its interaction energy written as:

$$E_{\mathrm{ts}}(\mathbf{r}) = \int |\phi(\mathbf{r}' - \mathbf{r})|^2 V_{\mathrm{ts}}(\mathbf{r}') d\mathbf{r}' \tag{6.5}$$

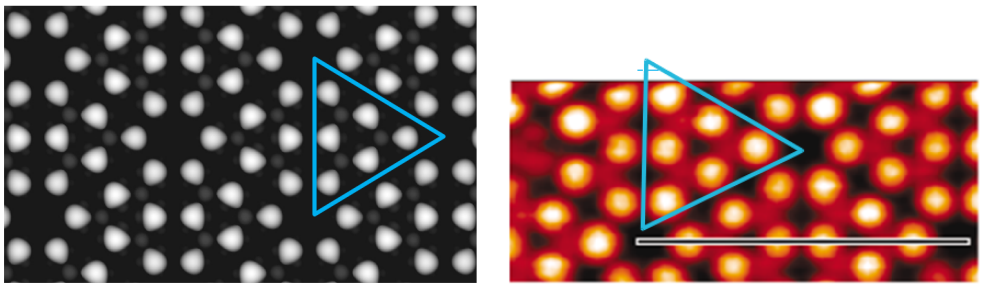
where V_{ts} is the potential on the tip from the sample and ϕ represents the electronic state of the tip. Equation 6.5 can be substituted into the force-

energy relation to give a convenient expression for the force on the tip:

$$\begin{aligned} F_{\mathrm{ts}}(\mathbf{r}) &= -\nabla E_{\mathrm{ts}}(\mathbf{r}) \\ &= -\nabla V_{\mathrm{ts}}(\mathbf{r}) - \nabla \left[\nabla V_{\mathrm{ts}}(\mathbf{r}) \int |\phi(\mathbf{r}' - \mathbf{r})|^2 (\mathbf{r}' - \mathbf{r}) d\mathbf{r}' \right] \\ &= -\nabla V_{\mathrm{ts}}(\mathbf{r}) - \nabla \left[\nabla V_{\mathrm{ts}}(\mathbf{r}) \cdot \mathbf{p} \right] \\ &= -\nabla V_{\mathrm{ts}}(\mathbf{r}) - \alpha \nabla \left(|\nabla V_{\mathrm{ts}}(\mathbf{r})|^2 \right) \\ &\propto -\nabla \left(|\nabla V_{\mathrm{ts}}(\mathbf{r})|^2 \right). \end{aligned} \tag{6.6}$$

The second line results from expanding V_{ts} up to the first order. In the following lines, $\mathbf{p}$ refers to the polarization of the tip, and α refers to its polarizability up to the first order. In the fourth line, the first term represents the monopole for electrons in the tip and is canceled out if the tip is electrically neutral. In experiment, this is done by applying a small tip-sample bias voltage. This leaves us with the second term, which represents the dipole interaction. If we assume the tip polarizability α is constant, then we obtain the final, simple expression for F_{ts} dependent only on the mapping of V_{ts}. In our electronic structure calculations, V_{ts} is constructed from the sum of the Hartree potential and the local component of the ionic pseudopotential. Therefore, using this expression for F_{ts} and the equations for the frequency shift given above, we can calculate an AFM image from a single self-consistent calculation.

Figure 6.3 shows a simulated AFM image with the virtual tip method for a Si(111)-(7x7) surface, which compares well with its experimental image. In practice, the virtual tip method has been shown to work favorably for surfaces or regular periodic lattices. It works less well for systems with drastic variations in the potential, such as in systems with defects or



(a) (b)

Figure 6.3: (a) Simulated AFM image of Si(111)-(77) surface with the virtual tip method [21]. (b) Experimental AFM image [33]. Triangles were drawn in for easier comparison.

in molecules with surface boundaries. Figure 6.4 shows AFM images for the dibenzo(cd,n)naphtho(3,2,1,8-pqra)perylene (DBNP) molecule. While the virtual tip simulated image is able to accurately reproduce the overall shape of the molecule, it cannot resolve the inner bond structure of the molecule compared to the full DFT simulated image with a CO tip. The resolution of the virtual tip is similar to that of the experimental image with the Xe-terminated tip. The difference in resolving power is ascribed to the special orbital characteristics of the CO molecule [62].

6.6 Tip tilting

Experimental and theoretical work have shown that tip tilting in the imaging process can have a noticeable effect on the AFM image [27, 43, 52, 45]. Methods for simulating tip tilting have been developed using classical potentials with dynamic relaxation [52, 48] as well as through density functional

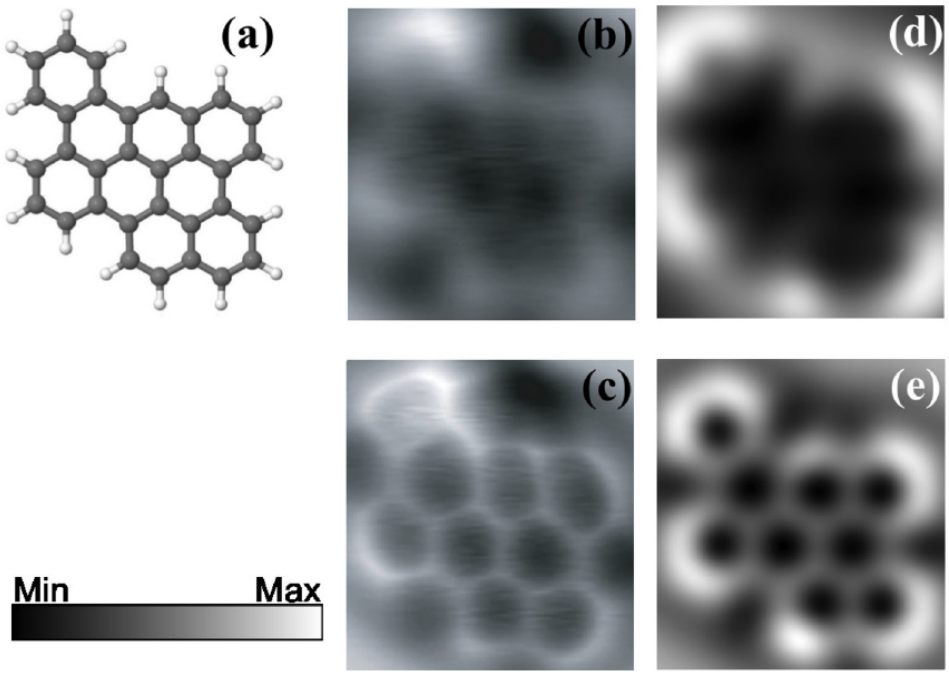


Figure 6.4: AFM images for dibenzo(cd,n)naphtho(3,2,1,8-pqra)perylene (DBNP) molecules. (a) Ball and stick model. Gray and white represent carbon and hydrogen atoms respectively. Experimental AFM images with (b) a Xe-terminated tip and (c) a CO-terminated tip [85]. PARSEC simulated images with (d) the virtual tip method and with (e) the full DFT method with a CO tip [62].

theory (DFT) calculations with "post-processing" [47].

To account for tip tilting we adopted the formalism developed by Guo *et al* [45, 47]. In this model, the tip probe (CO, for example) is assumed to be a rigid rod with a fixed bond length. When the tip is scanned over the sample, forces on the tip cause a lateral displacement of the downward pointing oxygen atom. The vertical displacement of the oxygen is assumed to be small and is thus neglected. The displaced position of the oxygen can be approximated as

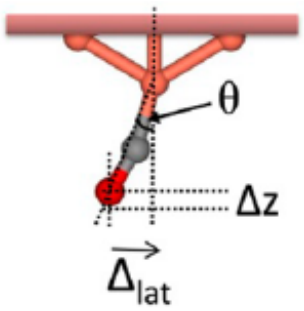
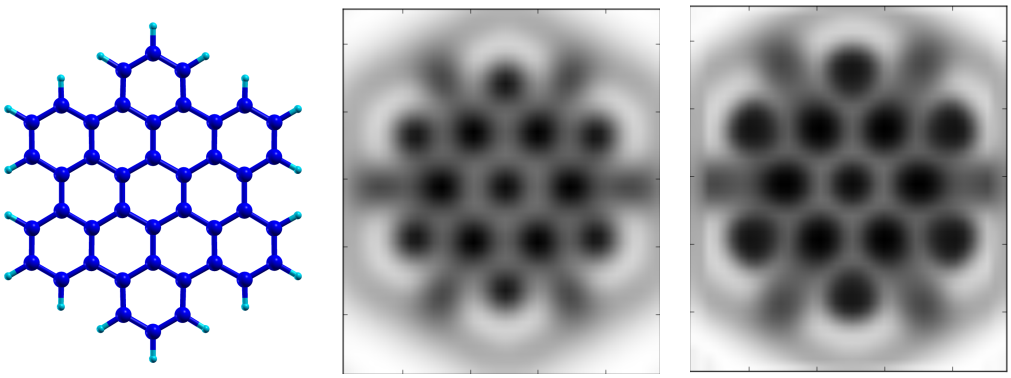


Figure 6.5: Tip tilting schematic [45].

a linear response to the lateral force by the equation $\vec{R}_{displaced} = \vec{R}_0 + \vec{F}_{lat}/k_{lat}$, where k_{lat} is a parameter for the lateral spring constant of the tip, and $\vec{F}_{lat}$ can be computed from numerical derivatives of the tip raster grid of interaction energies. The displaced positions are used with the tip raster grid to calculate tilt corrected values via cubic interpolation.



(a) (b) (c)

Figure 6.6: (a) Ball and stick model of hexabenzocoronene. Blue and cyan represent carbon and hydrogen atoms respectively. (b) Simulated AFM image at $z = 3.39$ Å. (c) Simulated image with tip tilting correction.

Figure 6.6 shows simulated AFM images of the simple hydrocarbon

molecule hexabenzocoronene with and without the tip tilting correction applied. We find that tip tilting tends to sharpen bonding features and distort images through expansion, generally making them look more like the experimental images. In some cases, tip tilting can resolve artifacts that can be mistaken for bonding features, which is a topic we will explore in the next chapter.

Chapter 7

Repulsive tip tilting as the dominant mechanism for hydrogen bond-like features

7.1 Work summary

Experimental atomic force microscopy (AFM) studies have reported distinct features in regions with little electron density for various organic systems. These unexpected features have been proposed to be a direct visualization of intermolecular hydrogen bonding. Here, we apply a computational method using *ab initio* real-space pseudopotentials along with a scheme to account for tip tilting to simulate AFM images of the 8-hydroxyquinoline dimer and related systems to develop an understanding of the imaging mechanism for hydrogen bonds. We find that contrast for the observed "hydrogen bond" feature comes not from the electrostatic character of the bonds themselves but rather from repulsive tip tilting induced by neighboring electron-rich atoms.

7.2 Introduction

Non-contact atomic force microscopy work by Zhang *et al.* has reported striking images of organic 8-hydroxyquinoline oligomers with distinct features between the moieties that have been interpreted as the hydrogen bonds that

hold the oligomers together [130]. However, the interpretation of these images is not well-understood, which has spurred numerous studies to examine the AFM imaging mechanism in detail [42, 46, 86, 43, 52, 53, 48, 111, 45, 47, 62].

The AFM imaging mechanism relies on an interaction between the probe tip and the sample being imaged. Experimental and theoretical work have shown that how the tip tilts in response to the sample during the imaging process can have a noticeable effect on the AFM image [27, 43, 52, 45]. Methods for simulating tip tilting have been developed using classical potentials with dynamic relaxation [52, 48] as well as through density functional theory (DFT) calculations with post-processing [47] to show that tip tilting can unexpectedly resolve features in regions of low electron density, apart from the typical covalent bonds of the molecule.

Our study uses *ab initio* pseudopotential-based density functional theory with a procedure to account for tip tilting to simulate nc-AFM images of the interacting 8-hydroxyquinoline dimer. We provide an interpretation for the measured image to examine the effect of the hydrogen bond on the contrast mechanism. We also examine a model system with a strong hydrogen bond: a hydrogen fluoride interacting dimer. In addition, we examine the organic molecule dibenzo[a,h]thianthrene (DBTH), a system that does not have a hydrogen bond but whose AFM image has shown features similar to one in experimental and computational studies [98, 52].

7.3 Computational methods

Our electronic structure calculations were carried out using PARSEC, a pseudopotential code for DFT calculations in real-space without the use of an explicit basis set [26, 69]. The real space aspect of PARSEC is well-suited for AFM simulations owing to PARSEC's flexible boundary conditions, which can eliminate the need for large, computationally intensive supercells to avoid artificial interactions due to imposed periodicity. For the systems in this study, we used localized boundary conditions in which the wave functions are required to vanish outside a large spherical domain. The size of this domain is constructed so that it has no influence on the solution to the electronic structure problem. We chose the size of the domain radius to be 24 a.u., 15 a.u., and 20 a.u. (1 a.u.$\sim$0.5292Å) for the 8-hydroxyquinoline, hydrogen fluoride, and DBTH systems, respectively. The Kohn-Sham problem was solved using high order finite differencing on a cubic grid with spacings of 0.25 a.u., 0.2 a.u., and 0.25 a.u. for the 8-hydroxyquinoline, hydrogen fluoride, and DBTH systems, respectively. We adopted Troullier-Martins norm-conserving pseudopotentials [119] and used the generalized gradient approximation of Perdew, Burke, and Ernzerhof (PBE) for the exchange-correlation functional [99]. Parameters for pseudopotentials were adopted from ABINIT's Fritz-Haber-Institute (FHI) pseudopotential database [38, 1]. Prior to performing AFM simulations, sample geometries were relaxed by the BFGS method [13, 36, 41, 106] with a force tolerance of 0.004 Ry/bohr ($\sim$ 0.1 eV/Å).

Using a simple model for the tip probe (a CO molecule), we construct a

2D raster grid of the tip over the sample and calculate interaction energies at different heights z to calculate the frequency shift. For the 8-hydroxyquinoline, hydrogen fluoride, and DBTH studies, we used tip raster grids of 0.5 a.u., 0.4 a.u., and 0.5 a.u., requiring 1872, 544, and 442 data points, respectively, for a given height. We used tip heights near 3.2 Å to obtain AFM images qualitatively comparable to experiment. To account for tip tilting we adopted the formalism developed by Guo *et al* [45, 47]. Our simulation method has been used to create accurate AFM images for organic systems such as pentacene and dibenzo(cd,n)naphtho(3,2,1,8-pqra)perylene [62].

We have considered the effects of using a more complex tip structure and including a substrate. Tests comparing the tip-sample interaction energy between Cu_2CO and CO found virtually no difference between the two. To test the substrate effect we mounted a sample of the 8-hydroxyquinoline dimer on four layers of Cu(111) in a periodic cell. We fixed the coordinates of the bottom two layers of the substrate, and we allowed the top two layers and the sample to relax. The resulting structure showed a slight corrugation, which had a very minor effect on the contrast of the simulated AFM images. Therefore, we used the CO molecule for the tip and excluded the substrate in our simulations. This procedure is justified by similar findings in other theoretical studies [46]. We note that the features shown in our images could not have come from the effect of complex tip geometries or the substrate since these were neglected in our simulations.

7.4 Results and discussion

We calculated AFM images for the 8-hydroxyquinoline dimer in the orientation with hydrogen bonds between the N and -OH atoms. The resulting image at a tip height of z= 3.15 Å (Fig. 7.1(a)) shows the expected features for the covalent bonds but no contrast between the moieties for the intermolecular bonds. After applying a tip tilting correction, the image (Fig. 7.1(b)) more closely resembles the experimental image (Fig. 7.1(c)), showing sharper bonds around the carbon rings and resolving the faint white lines between the moieties. However, an inspection of the atomic positions in our simulation shows that the hydrogen and nitrogen atoms are oriented vertically from each other on the y-axis, whereas the white lines of intermolecular contrast are slightly slanted. If one takes the hydrogen bond axis to be between electrostatic regions of partially positive and negative charge–the hydrogen and nitrogen atoms respectively–the intermolecular lines should also be vertical. In our image it almost appears as if the intermolecular line is between the oxygen and the nitrogen.

Various computational studies have been successful in resolving intermolecular features in AFM images through tip tilting with the dominant mechanism being thought to be Pauli repulsion between the tip and sample [52, 48, 45, 47]. To compare the effect of this mechanism with that of hydrogen bonding, we constructed a simple model system of the hydrogen fluoride dimer, which has a strong hydrogen bond. The calculated image (Fig. 7.2(a)) shows diffuse bright spots over the electron-rich fluorines. When

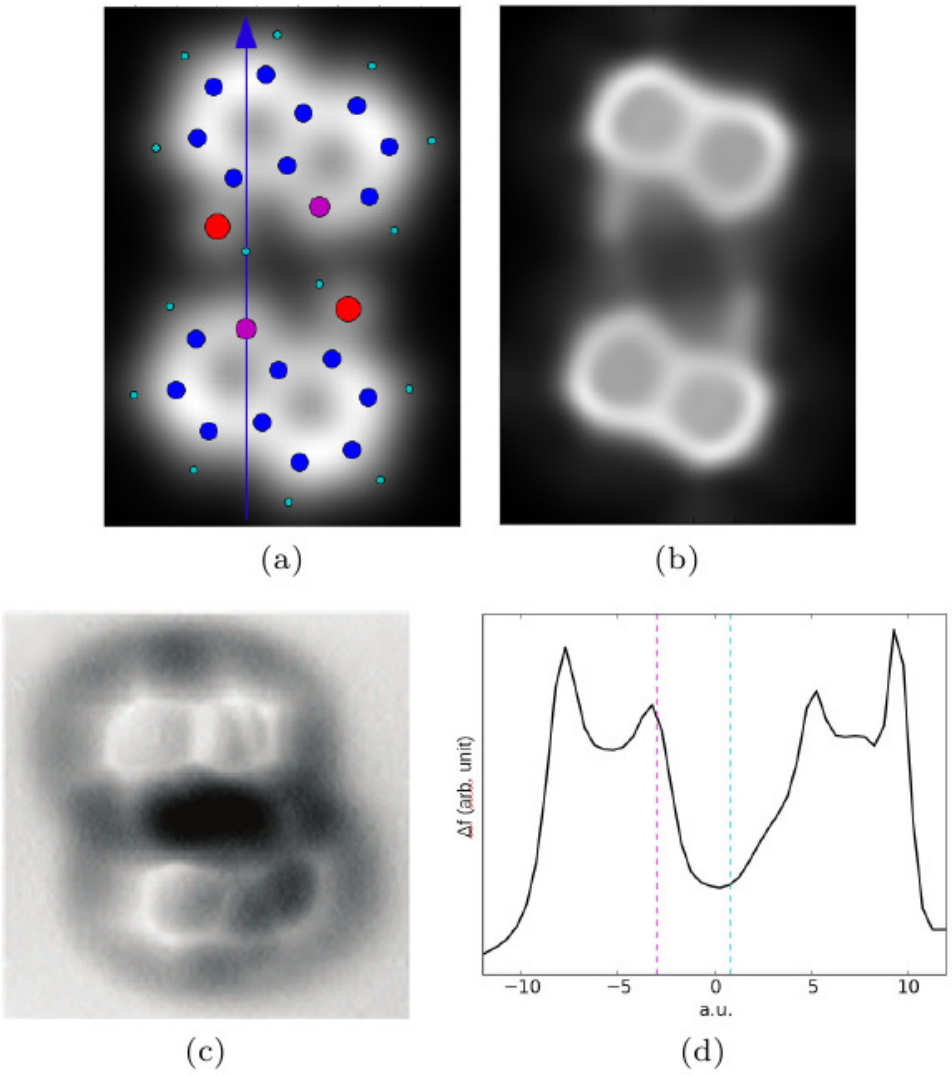


Figure 7.1: (a) Simulated AFM image of the 8-hydroxyquinoline dimer at $z = 3.18$ Å with atomic positions overlaid. Blue, red, magenta, and cyan represent carbon, oxygen, nitrogen, and hydrogen atoms respectively. (b) Image with tip tilting correction at $k_{lat} = 0.4$ N/m. (c) Experimental image from J. Zhang, P. Chen, B. Yuan, W. Ji, Z. Cheng, and X. Qiu, Science 342, 611 (2013). (d) Frequency shift of the tip tilted image along the line of the hydrogen bond axis as shown by the blue arrow in (a). The dashed magenta and cyan lines reflect the positions of the nitrogen and hydrogen atoms respectively.

the tip tilting correction is applied, these bright spots contract, and a faint "band" between the moieties starts to form (Fig. 7.2(b)). From this information alone, it is difficult to conclude that this band feature was caused by the hydrogen bond. Exploring this aspect in more detail, we compared this AFM image with that of the "flipped" orientation with the two fluorine atoms facing each other. Clearly, no hydrogen bond exists in this orientation as the partially negatively charged fluorines repel each other. It should be noted that the distance between the moieties was optimized separately in each orientation. While the distance between the moieties in the "flipped" orientation is longer than that of the hydrogen bonded orientation, the electron-rich fluorine atoms are closer to each other in the "flipped" case.

The resulting AFM images of the "flipped" orientation look strikingly similar to those of the hydrogen bonded orientation in both the regular and tip tilted cases (Fig. 7.2(c) and 7.2(d) respectively). We examine this feature more quantitatively by taking a line profile of the frequency shift along the central axis of the dimer (Fig. 7.2(e)). The line profiles for the hydrogen bonded and "flipped" orientations look nearly identical, with the hydrogen bonded profile showing a slight asymmetry in the region between the two fluorine peaks, which is emphasized in the tip tilted profile. The profiles show that the tendency of tip tilting is to sharpen the peaks of the large, electron-rich fluorine atoms. Interestingly, the magnitudes of the frequency shift at the maxima on top of the fluorine atoms and at the minima in between the fluorines are not affected by tip tilting; in fact, the frequency shift in the

intermolecular region of the tip tilted profile is lower than that of the original profile. Based on the profiles, an interpretation for the appearance of the intermolecular line follows. When tip tilting causes a sharpening of the peaks of strongly interacting atoms, the contrast in the broad tails of the peaks gets "swallowed up," and the intermolecular region is brought out by comparison, even though the frequency shift values are actually lower in the tip tilted case. In this sense, no phenomena in the intermolecular region is responsible for the appearance of the intermolecular line. Rather, it is the close proximity of electron-rich atoms through the effect of tip tilting that resolve the feature. This interpretation has been suggested by previous AFM simulation research using classical potentials [52, 48]. The line profile for the frequency shift across the intermolecular bond shows a similar peak-sharpening effect with the maxima occurring on the central axis of the dimer at $x = 0$.

We compared the line profiles of the frequency shift to that along the hydrogen bond axis in the 8-hydroxyquinoline dimer (Fig. 7.1(d)). This profile exhibits a large peak for the electron-rich nitrogen that decays along the length of the hydrogen bond, showing no feature over the hydrogen atom (the other large peaks in the profile represent the crossing of the carbon-carbon covalent bonds). If these profiles suggest that the dominating influence on the frequency shift image comes from electron-rich regions such as the fluorine and nitrogen atoms, an explanation for the slanted hydrogen bond lines in the 8-hydroxyquinoline dimer becomes possible. The intermolecular feature occurs not between the nitrogen and hydrogen of the hydrogen bond, but rather from

the nitrogen and the *neighboring oxygen* of the hydrogen atom in the hydrogen bond. If the hydrogen were to have a role in the formation of the hydrogen bond-like feature, it would be to polarize the neighboring oxygen, which could enhance tip tilting through repulsion from the electron-rich oxygen and nitrogen atoms. The electrostatic character of the hydrogen bond itself is thought to have little effect on the AFM image.

As an additional computational "experiment," we calculated an AFM image for dibenzo[a,h]thianthrene (DBTH) in the folded configuration with upward-pointing sulfur atoms (Fig. 7.3(a)). This molecule is of special interest in AFM theory because experimental images resolve a white line between the two para-sulfurs on the central carbon ring where no chemical bond exists (Fig. 7.3(d)). The folding angle was optimized at around 136°, close to the previous theoretical value of 133° [97]. The calculated AFM images (Fig. 7.3(b) and 7.3(d)) look similar to those of the HF dimer, where bright diffuse spots appear for the electron rich sulfur atoms. When tip tilting is accounted for, these bright spots condense and a distinct feature forms between the two sulfurs. The imaging mechanism should be the same as for the hydrogen fluoride and 8-hydroxyquinoline dimers. The carbons do not appear on the image because they are bent away too far away from the tip probe. Simulations with a dynamic method for tip relaxation using classical potentials have also showed this result [52], although it should be noted that the DFT simulated image matches with experiment more closely.

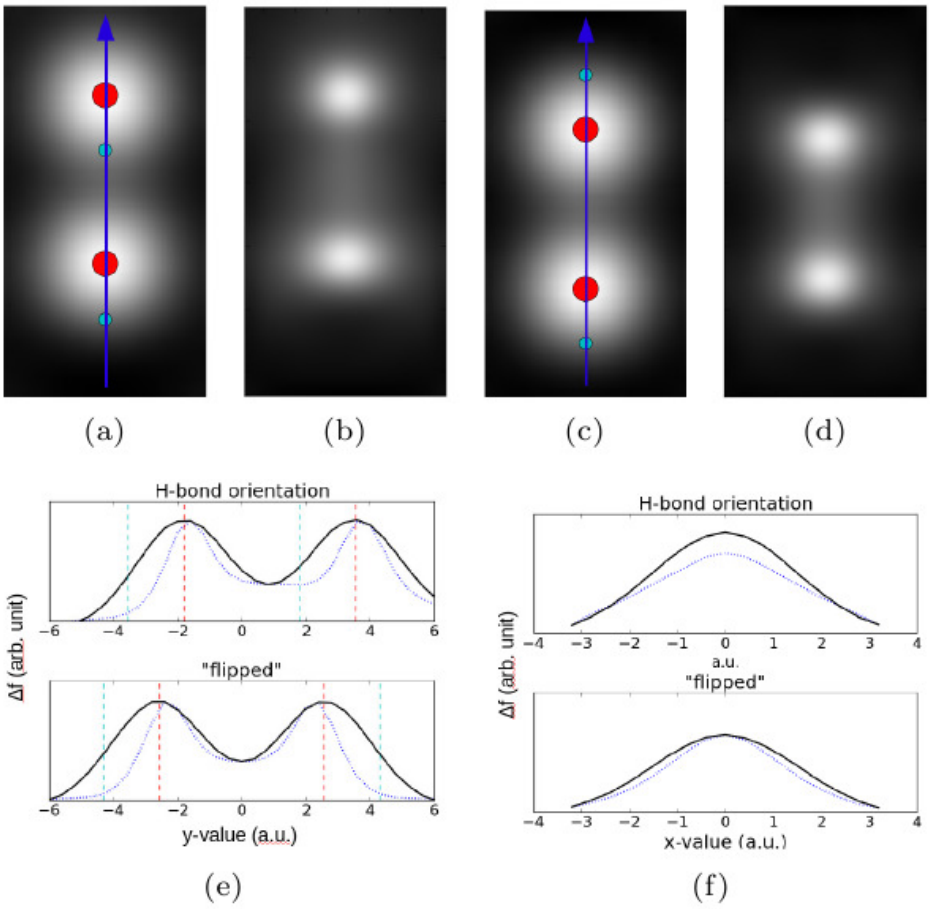


Figure 7.2: (a) Simulated AFM image of a hydrogen fluoride dimer with a hydrogen bond at $z = 3.18$ Åwith atomic positions overlaid. Red and cyan represent fluorine and hydrogen atoms respectively. (b) Image with tip tilting correction. (c)-(d) Same as (a)-(b) except for the "flipped" orientation. (e) Frequency shift of the original (black) and tip tilted (dotted blue) image *along* the central axis of the dimer (x = 0) as shown by the blue arrow in (a) and (c). The dashed red and cyan lines reflect the positions of the fluorine and hydrogen atoms respectively. (f) Same as (d) except *across* the intermolecular region at $y = 0$.

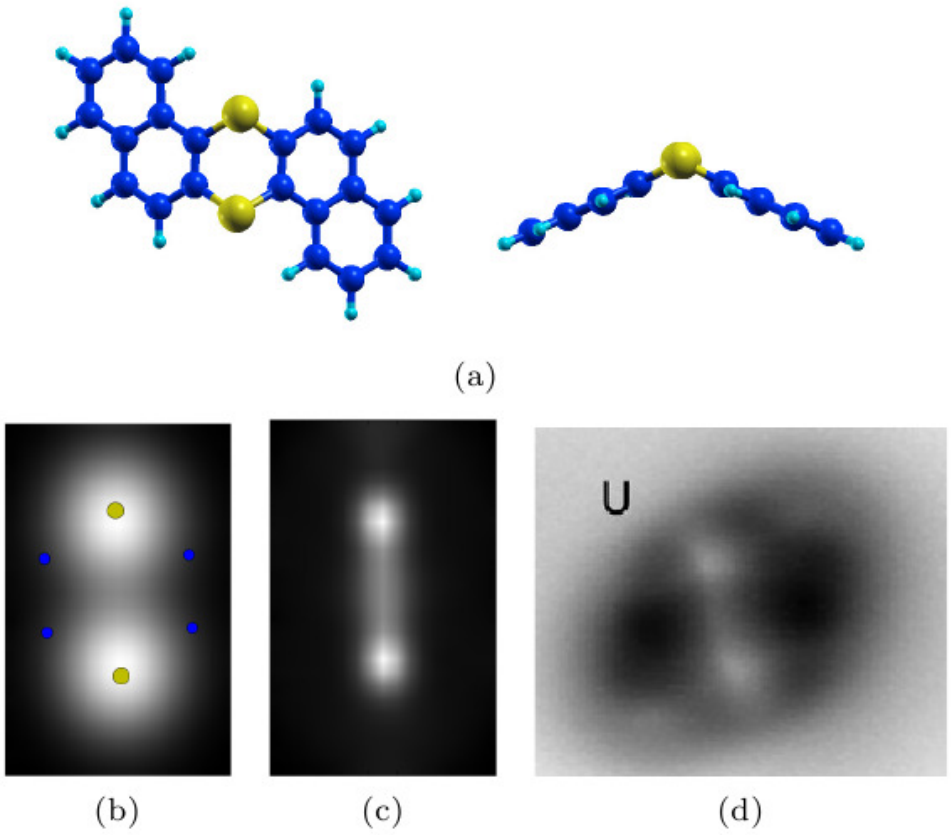


Figure 7.3: (a) Top (left) and side (right) views of DBTH with upward-pointing sulfurs (U). The blue, yellow, and cyan atoms represent carbon, sulfur, and hydrogen respectively. Molecular structure images were creating using XCrySDen [66]. (b) Simulated AFM image at z = 3.37 Åwith atomic positions overlaid. (c) Image with tip tilting correction. (d) Experimental image of DBTH from N. Pavliček, C. Herranz-Lancho, B. Fleury, M. Neu, J. Niedenführ, M. Ruben, and J. Repp, Phys. Status Solidi B 250, 2424 (2013).

7.5 Conclusions

AFM simulations have been able to reproduce the image of the 8-hydroxyquinoline dimer and its hydrogen bond. However, these simulations suggest that the intermolecular feature comes not from a direct imaging of the hydrogen bonds themselves. The hydrogen atoms in the hydrogen bonds polarize the neighboring oxygen atoms, and repulsive tip tilting resolves hydrogen bond-like lines between the oxygen and nitrogen atoms. The electrostatic nature of the hydrogen bond itself is thought to have little effect on the imaging mechanism. A better understanding of the effect of intermolecular contrast brought out by tip repulsion would aid the interpretation of experimental images. For example, a more quantitative approach could examine the effects on tip tilting from the intermolecular bond distance, the character of electron-rich atoms, and polarization.

Chapter 8

Contrast inversion induced by chemically reactive tips

8.1 Work summary

Experimental atomic force microscopy (AFM) studies have reported contrast inversions for systems such as Cu_2N and graphene that can hamper image interpretation and characterization. Here, we apply a simulation method based on *ab initio* real-space pseudopotentials in order to gain an understanding of the tip-sample interactions that influence the inversion. We find that reactive metal tips such as Cu_2 induce an attractive binding force that brings about the contrast inversion. The inversion is tip height dependent and not observed when using chemically inert tips such as CO. The contrast inversion cannot be accurately modeled with simulation methods that assume a weak interaction between tip and sample.

8.2 Introduction

One particular phenomenon found in AFM imaging studies is "contrast inversion." In an experimental study, graphene imaged with a CO-functionalized tip shows brightness along the carbon-carbon covalent bonds and dark spots in

the centers of the carbon rings where no bonding occurs [8]. This is a typical contrast pattern for an AFM image, where brightness appears around bonding and atomic sites. However, when a metal iridium tip is used as the probe, at tip heights close to the sample the contrast inverts such that the covalent bonds show up as dark regions and the centers of the rings show up as bright spots. Understanding what causes contrast inversion and knowing when it occurs is of crucial importance for interpreting images accurately. If the inversion is not taken into account, an interpretation can be directly opposite of what the image actually shows!

Contrast inversions have also appeared in AFM imaging studies on metal Cu_2N surfaces. To examine this problem, a research group developed a simple AFM simulation method based purely on the electrostatic interaction between the tip and sample [105]. This electrostatic model replicated the inversion characteristics of experimental images of Cu_2N for two different probes, metal Cu and CO. In their discussion, the primary factor is not the reactivity of the tip but rather its dipole. The metal Cu tip has a dipole with its negative end pointing away from the sample whereas the CO tip has it pointing toward the sample. This difference is thought to contribute to the inversion characteristics of the experimental images.

The main principle behind AFM imaging is the interaction between the tip probe and the sample being imaged. As such, the chemical makeup of the tip can have drastic effects on the resulting image. The previous studies suggest that multiple mechanisms are responsible for contrast inversions. One study

suggests that the property responsible for inversion is the chemical reactivity of the tip (reactive metal Ir vs. non-reactive CO). The other suggests that the key factor is the electrostatic dipole of the tip. We will use simulations to probe these behaviors and see how they contribute to the final image.

Theoretical AFM simulations can be difficult to carry out because the morphology of the probe tip and the dynamics during the scanning process are not precisely known. Various models, which often involve the appropriate use of simplifying assumptions, have been developed to make this problem computationally tractable. For example, the electrostatic model used in the aforementioned Cu_2N study neglects short-range Pauli forces, which are thought to play an important role in the AFM imaging mechanism at close tip distances. Another electrostatic model called the "virtual tip" method avoids the explicit modeling of the tip [21]. Simulation methods that use classical potentials [52, 48] or *ab initio* quantum calculations [47, 62, 73, 95] to describe the tip-sample interaction are not limited to electrostatics but pay the price in computational cost. Knowing the appropriate model to select for a specific system can result in the optimal combination of accuracy and computational efficiency.

In this paper, we will use an AFM simulation method based on first-principles density functional theory (DFT) calculations in order to study tip reactivity and its role in image contrast inversion. We will look at two systems, graphene and a metal Cu_2N surface, each with two different tips: a reactive metal (Cu) tip and a chemically inert CO tip. We will also apply a

simulation method based on frozen density embedding theory (FDE) theory, which is an approximation devised to speed up the cost of performing repeated DFT calculations. By comparing results from the different models and with experimental images, we clarify the mechanism for contrast inversion in AFM imaging studies.

8.3 Computational Methods

Electronic structure calculations were carried out using PARSEC, a pseudopotential code for DFT calculations in real-space without the use of an explicit basis set [26, 69]. For both the graphene and Cu_2N systems, we used slab (2D) boundary conditions, which imposes periodicity in the lateral x and y directions. Along the z-axis, the wave functions are required to vanish outside a defined domain height. We chose a sufficiently large height so that the boundary has no influence on the solution to the electronic structure problem. This required around 10 a.u. of vacuum space between the outermost atom and the edge of the domain. (1 a.u.$\sim$0.5292Å)

To construct the Cu_2N system, we placed a Cu_2N monolayer on top of a Cu(100) surface in a 2 $\times$ 2 supercell (Fig. 8.1(a)). For the full DFT simulations, we included two layers of Cu(100) below the monolayer and six layers for the FDE simulation. To construct the graphene system, we used a $\sim$18.41 $\times$ 15.95 a.u. rectangular supercell and did not include an underlying substrate. Modeling the underlying substrate in AFM simulations can be challenging because it not only influences the periodicity of the system but

also greatly increases the size of the computational problem. In a previous study, we tested the effect of including a Cu substrate for pentacene and found no significant contribution to the AFM image, so we left it out for the purpose of computational efficiency [104].

The Kohn-Sham problem was solved using high order finite differencing with a grid spacing of 0.3 a.u. and a k-point mesh of 3×3. We adopted Troullier-Martins norm-conversing pseudopotentials [119] and used the local density approximation (LDA) constructed by Ceperley-Alder [17] for the exchange-correlation functional. Copper pseudopotentials were constructed based on a previous paper [107]. Prior to performing AFM simulations, the geometries of the samples were relaxed by the BFGS minimization algorithm [13, 36, 41, 106] with a force tolerance of 0.004 Ry/a.u. ($\sim$ 0.1 eV/Å). For the Cu_2N calculations, we used a Fermi temperature of 800 K to help with the total energy convergence.

Frozen density embedding theory (FDE) can be applied to AFM imaging simulations as an approximation to full DFT calculations. We used the kinetic energy functional proposed by Tran and Weslowszki (PBE-TW) [118]. One of the key assumptions of the FDE method is that the interaction between the tip and sample is small enough that the tip does not affect the structural or electronic properties of the sample. By comparing the results of the FDE simulations with the full DFT ones, we will be able to test how well these assumptions hold. We calculated the frequency shift using the small amplitude approximation. With the help of FDE, we tested the approximation and

found no notable difference between images computed by Eq. 6.1 or 6.2.

We model the tip probe as a simple two atom system (Cu_2 for the reactive metal tip and CO for the chemically inert tip). We then construct a 2D raster grid of the tip over a region of the sample and calculate interaction energies at different heights z to calculate the frequency shift using Eq. 6.2. We found that using raster grids of twice the grid spacing (0.6 a.u.) provided sufficient resolution for our AFM images. For a given tip height z, the number of data points computed were 324 and 420 for the Cu_2N and graphene systems, respectively.

To test the influence of the tip on substrate relaxation, we placed the reactive Cu_2 tip above graphene at a close distance ($\sim$2.1 Å). We fixed the tip at various sites and allowed the sample to relax around it. The sample atoms hardly moved from their original positions. Therefore, it is expected that substrate relaxation at the larger tip heights used in the study should not affect the AFM images. It is also possible for the AFM image to be affected by the tip tilting in response to forces from the sample. Since we did not include a scheme for tip tilting, the inversion characteristics in the study can not be accounted for by tip tilting.

8.4 Results and discussion

First we simulated AFM images of Cu_2N with the reactive metal Cu_2 tip (Fig. 8.1(b) and the chemically inert CO tip (Fig. 8.1(c)). The two tips do not give the same contrast mapping. For the Cu_2 tip image, the brightest

spots occur above the hollow sites, and the darkest regions occur above the nitrogen sites. In the other case, the CO tip image is a near inversion of the Cu_2 tip image. Here the hollow sites show the darkest contrast whereas the brightest spots occur above the nitrogen atoms. These images match previous experimental findings on Cu_2N with Cu- and CO- terminated tips [105].

To examine the contrast mapping more closely, we plotted the vertical force on each tip as a function of height over the different sites (Fig. 8.2). Comparing the results for the two tips, the force curves are strikingly different. With the Cu_2 tip, the forces become more negative (attractive) with decreasing tip height. Across most of the scanned region, the order of forces from lowest to highest is the nitrogen site followed by the copper site and finally the hollow site. This corresponds to the nitrogen sites being the darkest and the hollow sites being the brightest in the AFM image. For the CO tip, the forces trend positive (repulsive) with decreasing tip height, and the order of forces is reversed. In AFM measurements, the balance between forces is typically between short-range attractive and Pauli repulsive forces while long-range van der Waals forces supply an attractive background that sets the scale of the forces [86, 95, 34]. With the reactive Cu_2 tip, the short-range attractive force dominates such that the forces above the atomic sites drop below that of the hollow site. In the case of the chemically inert CO tip, the low reactivity of the tip causes the repulsive force to be dominant, so the forces above the atomic sites now become larger than those of the hollow site.

To visualize the difference in the reactivities of the Cu_2 and CO tip, we

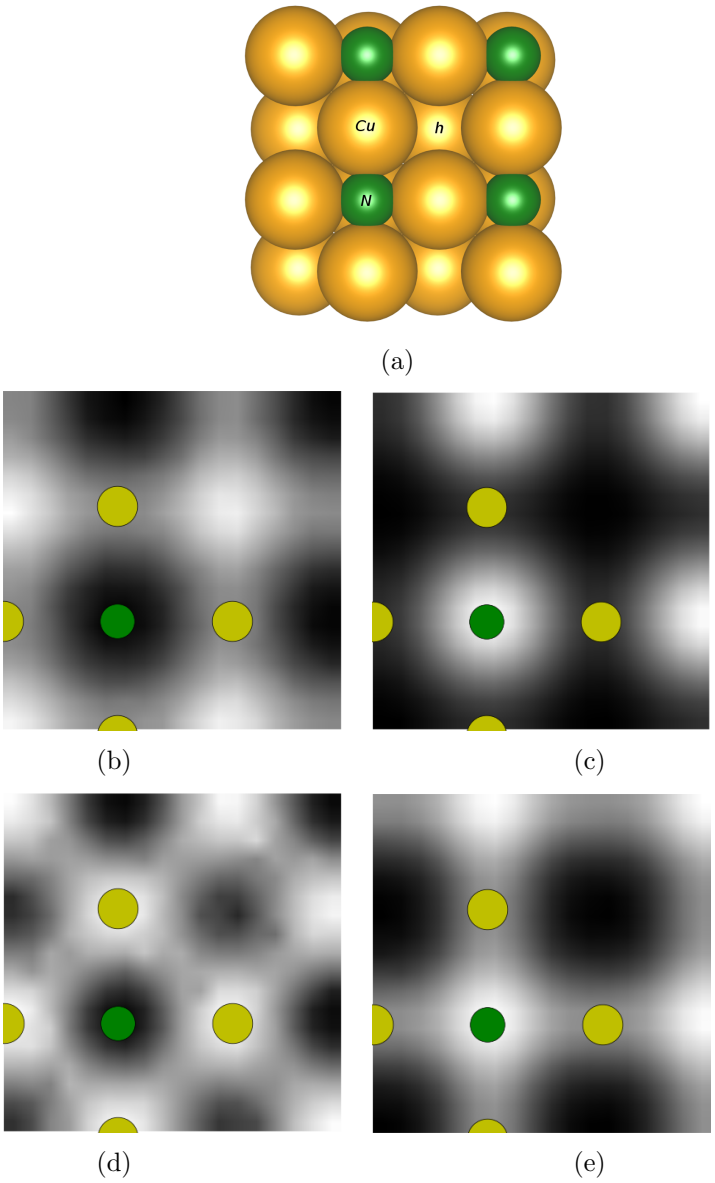


Figure 8.1: (a) Top view of a 2 x 2 supercell of Cu_2N over Cu(100). Yellow and green atoms represent copper and nitrogen, respectively. Cu, N, and h labels designate the copper site, nitrogen site, and hollow site, respectively. Molecular structure image was created using VESTA [87]. Full DFT simulated AFM images of Cu_2N with (b) Cu_2 tip at $z = 3.55$ Å and with (c) CO tip at $z = 3.07$ Å. Simulated images using FDE with (d) Cu_2 tip at $z = 3.70$ Å and with (e) CO tip at $z = 3.39$ Å. Atomic positions are partially overlaid.

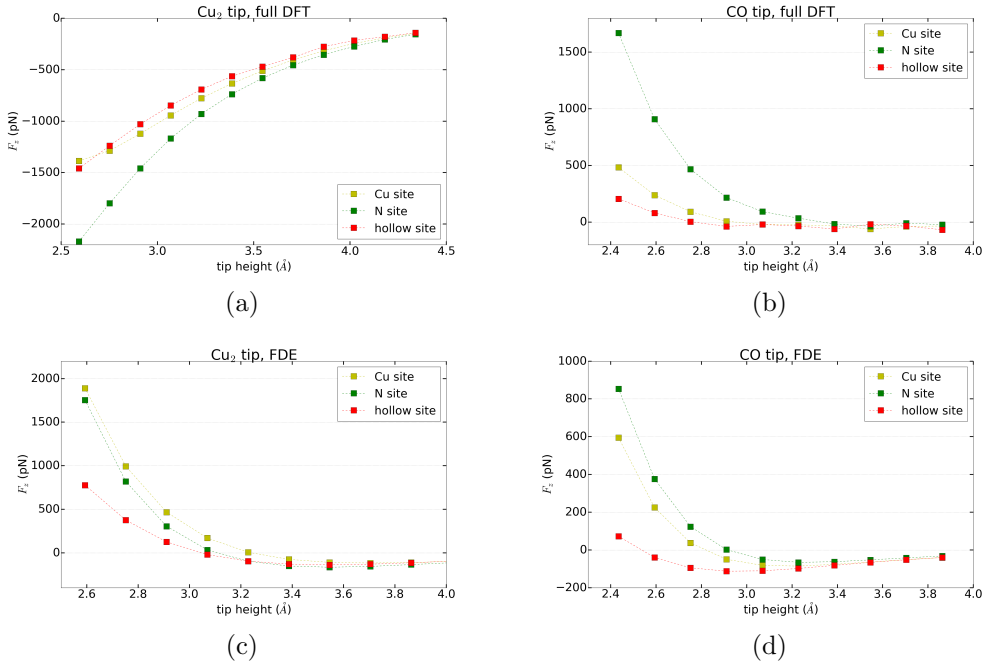


Figure 8.2: (a)-(d) Vertical force as a function of tip height above various sites on Cu_2N. See Fig. 8.1(a) for labeling of the sites, and refer to individual plots for tip selection (Cu_2 or CO) and simulation method (full DFT or FDE).

calculated charge density difference plots from the equation:

$$\rho_{diff} = \rho_{tip+substrate} - \rho_{tip} - \rho_{substrate} \tag{8.1}$$

where the charge densities of systems with only the tip and only the substrate are subtracted from that of the combined tip and substrate system (Fig. 8.3). We select the value of the isosurface to be the same across the plots (3×10^{-4} electrons/a.u.^{3}) in order to make quantitative comparisons. The plots for the Cu_2 tip (Fig. 8.3(a)-8.3(c)) clearly show a strong increase in electron density between the tip and sample whereas for the CO tip (Fig. 8.3(d)-8.3(f)), hardly any change in electron density is observable, even though the tip distance is closer to the sample compared to that of the Cu_2 tip case.

As we have shown that strong binding between the tip and sample is a large influence on the contrast mapping for the Cu_2 tip, we now test if frozen density embedding theory can be used to accurately simulate systems with reactive tips. FDE simulated images for the Cu_2 and CO tips are shown in Fig. 8.1(d)-8.1(e). The FDE image with the reactive Cu_2 tip does not match the full DFT image. While both images have the darkest region at the nitrogen site, the FDE image has the copper site as its brightest spot rather than the hollow site. The FDE force curve (Fig. 8.2(c)) conspicuously fails to capture the short-range attractive forces on the tip as the forces are repulsive, having the incorrect sign. On the other hand, the FDE image with the chemically inert CO tip fares much better and matches well with the full DFT case, with the noticeable difference of the copper sites being much brighter than in the full

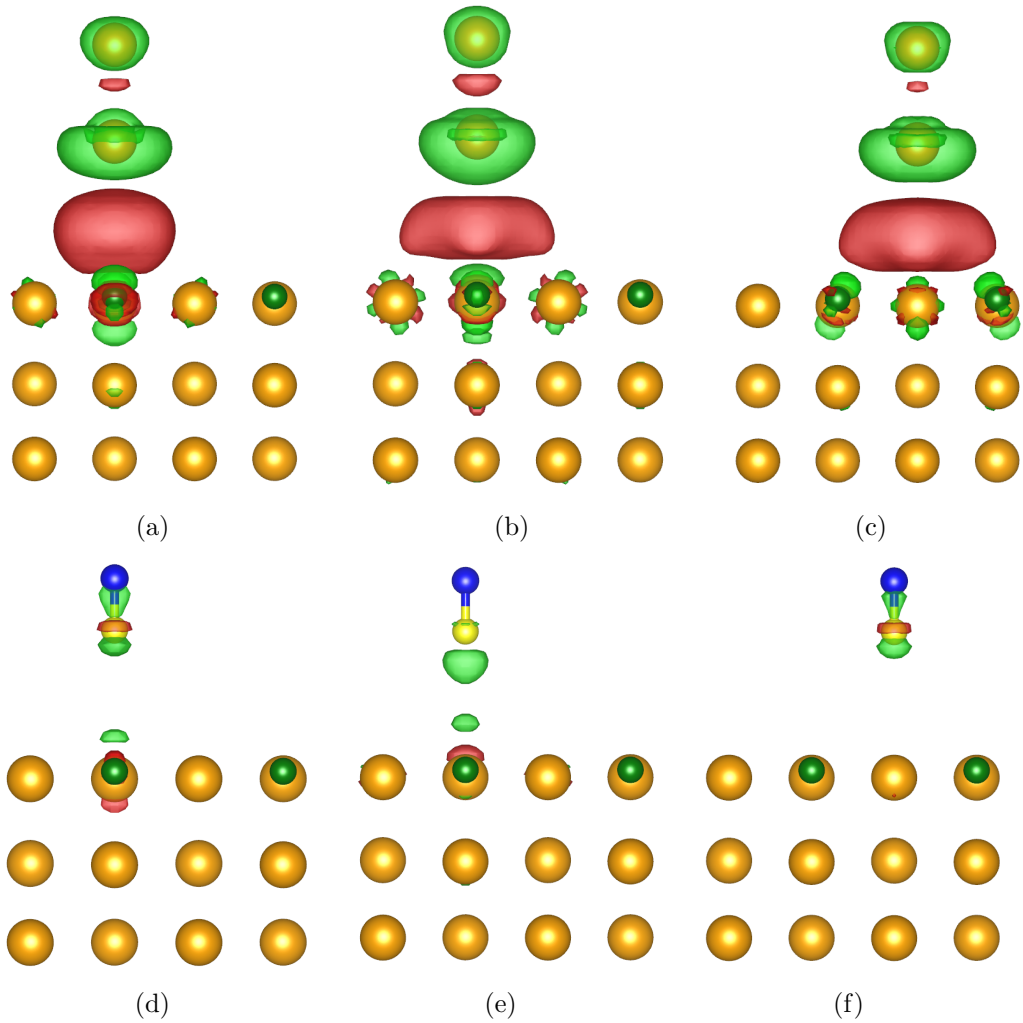


Figure 8.3: Charge density difference plots for Cu_2N with Cu_2 tip at $z = 3.55$ Åabove (a) copper site, (b) nitrogen site, and (c) hollow site. Same as previous except with CO tip at $z = 3.07$ Å above (d) copper site, (e) nitrogen site, and (f) hollow site. The blue and yellow spheres of the tip represent carbon and oxygen atoms respectively. Red and green represent positive and negative charge density respectively. For both sets of plots, the isosurface was set to the same value (3×10^{-4} electrons/a.u.^{3}) for quantitative comparison. See Fig. 8.1(a) for labeling of the sites. Images were created using VESTA [87].

DFT case. Comparing the FDE and full DFT force curves shows reasonable agreement with a tip height offset of about 0.3 Å, which is similar to that in a previous study [104].

We repeated the AFM image analysis with the reactive and chemically inert tips on graphene (Fig. 8.4). At tip heights of around 3 Å, the Cu_2 and CO tip images are inverted from each other with respect to the contrast on the carbon and hollow sites. However, if the Cu_2 tip is brought close enough to the graphene surface, it inverts back to looking like the CO tip image, with brightness along the regions of the carbon-carbon bonds. The CO tip image does not invert when the tip is brought closer to the sample. These observations are consistent with experimental measurements of graphene using a reactive metal (Ir) and chemically inert (CO) tip [8].

The force curves (Fig. 8.5) clearly show how the contrast inverts with tip height. For the Cu_2 tip case, the force curves above the two sites cross at around 2.2 Å whereas for the CO tip case the force curves never cross. The shape of the force curves can be explained by the reactivity of the tip. To visualize this, we plotted the charge density differences ρ_{diff} above the carbon and hollow sites for the Cu_2 tip at the height $z = 3.07$ Å (Fig. 8.6). For both plots we choose the same value for the isosurface (9×10^{-4} electrons/a.u.^{3}). At this intermediate distance, the tip binds more strongly to electronic states above the carbon site compared to the hollow site, lowering the force above the carbon site and thus creating the inversion. At close tip height distances, repulsive forces overtakes the attractive force and causes a steep rise in the

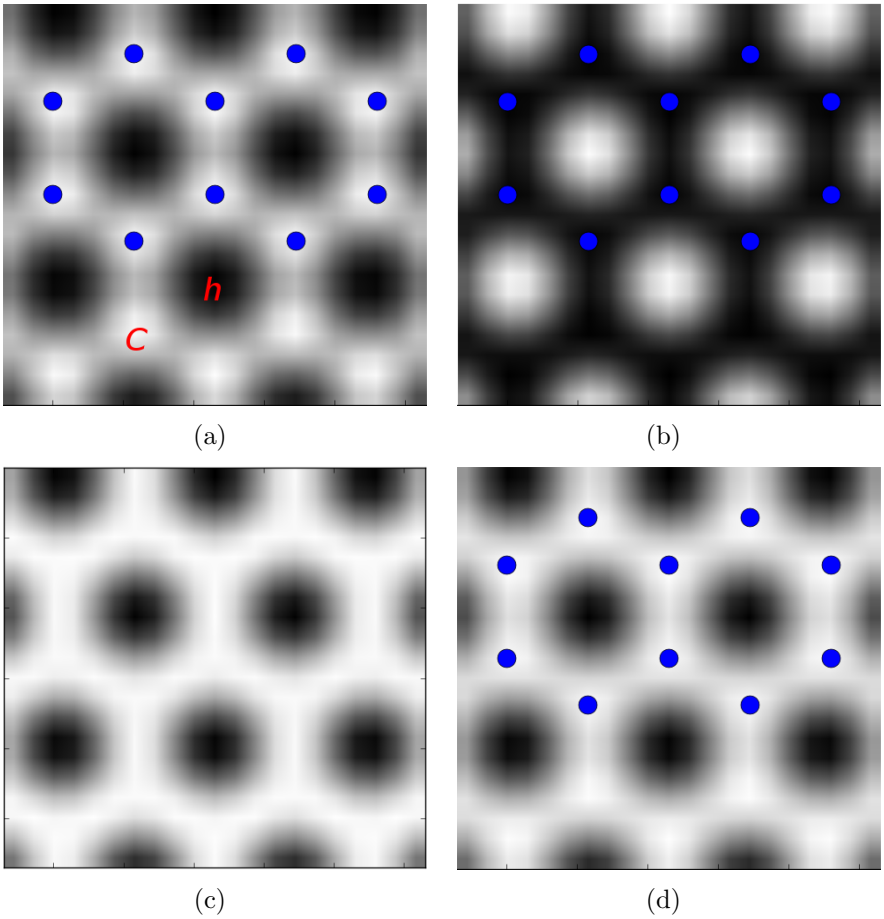


Figure 8.4: Full DFT simulated AFM images of graphene with Cu_2 tip at (a) $z = 2.11$ Å and at (b) $z = 3.07$ Å. C and h labels designate the carbon site and hollow site respectively. Image with CO tip at (c) $z = 2.11$ Å and at (d) $z = 3.07$ Å. Atomic positions are partially overlaid.

force above the carbon site, and the contrast inverts back to match that of the CO tip image.

Finally, we tested the FDE imaging method on graphene with both tips. Once again, we find that FDE fails to capture the correct inversion characteristics of the reactive Cu_2 tip, as the force curves shows no crossing (Fig. 8.5(c)). For the chemically inert tip, the FDE measurement shows good agreement with the full DFT case after accounting for the usual offset of about 0.3 Å. These results suggest that FDE would be the ideal method for simulating graphene with chemically inert tips such as CO because the computational gains from the method can be obtained without loss of accuracy. However, for situations where tip reactivity is thought to be an important factor, an alternative simulation method would be preferable.

8.5 Conclusions

We simulated AFM images of Cu_2N and graphene systems with two different tip types: a reactive metal Cu_2 tip and a chemically inert CO tip. The reactive tip induces a binding attraction that lowers the force over atomic sites such that the image contrast is inverted compared to the chemically inert tip image. This contrast inversion is tip height dependent. As the tip moves closer to the sample, steeply rising repulsive forces overtake the attractive forces over the atomic sites and induce the image contrast to reinvert. The study highlights the importance of tip selection, tip height, and simulation method when performing AFM experiments, as different methods influence

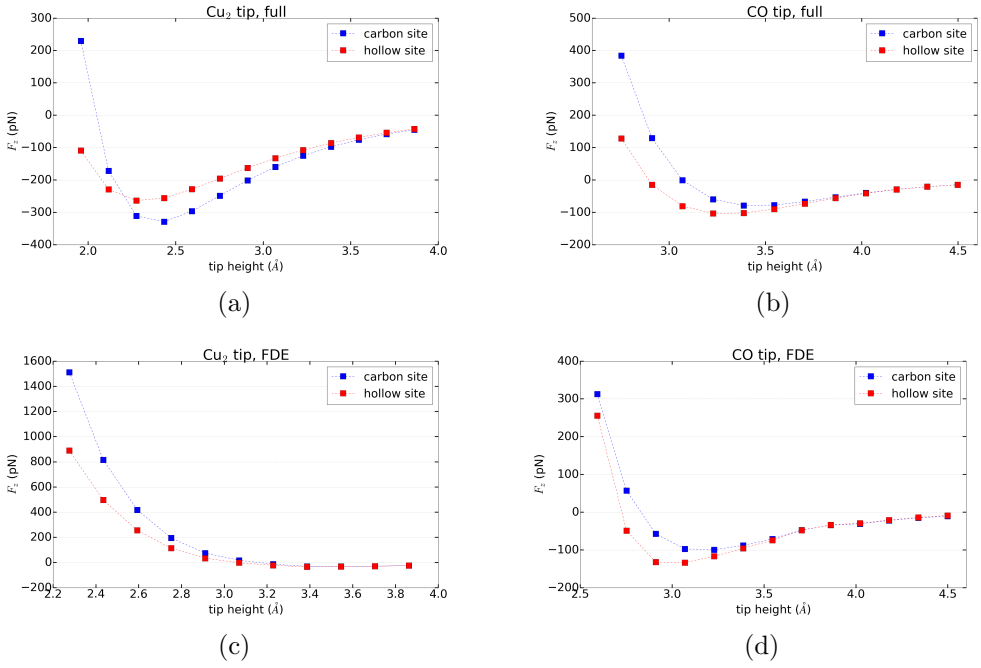
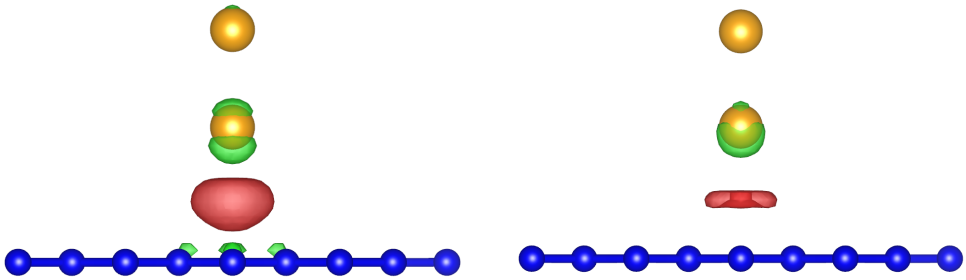


Figure 8.5: (a)-(d) Vertical force as a function of tip height above various sites on graphene. See Fig. 8.4(a) for labeling of the sites, and refer to individual plots for tip selection (Cu_2 or CO) and simulation method (full DFT or FDE).



(a) (b)

Figure 8.6: Charge density difference plots for graphene with Cu_2 tip at z = 3.07 Å above the (a) carbon site and (b) hollow site. Red and green represent positive and negative charge density respectively. For both plots, the isosurface was set to the same value (9×10^{-4} electrons/a.u.^{3}) for quantitative comparison. See Fig. 8.4(a) for labeling of the sites. Images were created using VESTA [87].

the types of forces on the tip central to the AFM imaging process.

Bibliography

[1] `http://www.abinit.org/downloads/psp-links/pseudopotentials`.

[2] M. M. G. Alemany, M. Jain, L. Kronik, and J. R. Chelikowsky. *Phys. Rev. B*, 69:075101, 2004.

[3] J. T. Arantes and A. Fazzio. *Nanotechnology*, 18:295706, 2007.

[4] G. Audoit, E. N. Mhuircheartaigh, S. M. Lipson, M. A. Morris, W. J. Blau, and J. D. Holmes. *J. Mater. Chem.*, 15:4809, 2005.

[5] S. P. Beckman, J. Han, and J. R. Chelikowsky. *Phys. Rev. B*, 74:165314, 2006.

[6] M. Bescond, N. Cavassilas, K. Nehari, and M. Lannoo. *J. Comp. Elect.*, 6:341, 2007.

[7] M. T. Björk, H. Schmid, J. Knoch, H. Riel, and W. Riess. *Nat. Nanotechnol.*, 4:103, 2009.

[8] M. P. Boneschanscher, J. van der Lit, Z. X. Sun, I. Swart, P. Liljeroth, and D. Vanmaekelbergh. *ACS Nano*, 6:10216, 2012.

[9] E. Borchi, S. D. Gennaro, R. Macii, and M. Zoli. *J. Phys. D: Appl. Phys.*, 21:1304, 1988.

[10] M. Born and R. Oppenheimer. *Ann. Physik*, 389:457, 1927.

[11] A. I. Boukai, Y. Bunimovich, J. Tahir-Kheli, J-K. Yu, W. A. Goddard, and J. R. Heath. *Nature*, 451:168, 2008.

[12] W. R. Bowen and N. Hilal. *Atomic Force Microscopy in Process Engineering*. Elsevier, 2009.

[13] C. G. Broyden. *IMA J. Appl. Math.*, 6:222, 1970.

[14] F. Bruneval. *Phys. Rev. Lett.*, 103:176403, 2009.

[15] W. R. Burdick, Y. Saad, L. Kronik, I. Vasiliev, M. Jain, and J. R. Chelikowsky. *Comp. Phys. Comm.*, 156:22, 2003.

[16] D. M. Ceperley and B. J. Alder. *Phys. Rev. Lett.*, 45(7):577, 1980.

[17] D. M. Ceperley and B. J. Alder. *Phys. Rev. Lett.*, 45:566, 1980.

[18] T.-L. Chan. *Phys. Rev. B*, 86:245414, 2012.

[19] T-L. Chan, A. J. Lee, and J. R. Chelikowsky. *Phys. Rev. B*, 91:235445, 2015.

[20] T.-L. Chan, M. L. Tiago, E. Kaxiras, and J. R. Chelikowsky. *Nano Lett.*, 8 (2):596, 2008.

[21] T.-L. Chan, C. Z. Wang, K. M. Ho, and J. R. Chelikowsky. *Phys. Rev. Lett.*, 102:176101, 2009.

[22] T.-L. Chan, S. B. Zhang, and J. R. Chelikowsky. *Phys. Rev. B*, 83:245440, 2011.

[23] T.-L. Chan, S. B. Zhang, and J. R. Chelikowsky. *App. Phys. Lett.*, 98:133116, 2011.

[24] J. R. Chelikowsky. *J. Phys. D*, 33:R33, 2000.

[25] J. R. Chelikowsky, M. M. G. Alemany, T.-L. Chan, and G. M. Dalpian. Reports on progress in physics volume: 74 issue: 4. *Rep. Prog. Phys.*, 74:046501, 2011.

[26] J. R. Chelikowsky, N. Troullier, and Y. Saad. *Phys. Rev. Lett.*, 72:1240, 1994.

[27] C-L. Chiang, C. Xu, Z. Han, and W. Ho. *Science*, 344:885, 2014.

[28] G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán. *Phys. Rev. B*, 79:155107, 2009.

[29] Y. Cui, Q. Wei, P. Park, and C. M. Lieber. *Science*, 293:1289, 2001.

[30] A. G. Cullis, L. T. Canham, and P. D. J. Calcott. *J. Appl. Phys.*, 82:909, 1997.

[31] G. M. Dalpian and J. R. Chelikowsky. *Phys. Rev. Lett.*, 96:226802, 2006.

[32] M. Diarra, Y. M. Niquet, C. Delerue, and G. Allan. *Phys. Rev. B*, 75:045301, 2007.

[33] T. Eguchi and Y. Hasegawa. *Phys. Rev. Lett.*, 89:266105, 2002.

[34] M. Ellner, N. Pavlicek, P. Pou, B. Schuler, N. Moll, G. Meyer, L. Gross, and R. Perez. *Nano Lett.*, 16:1974, 2016.

[35] P. P. Ewald. *Ann. Physik*, 64:253, 1921.

[36] R. Fletcher. *Comp. J.*, 13:317, 1970.

[37] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle. *Phys. Rev. Lett.*, 102:016402, 2009.

[38] M. Fuchs and M. Scheffler. *Comp. Phys. Comm.*, 119:67, 1999.

[39] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. Fabris, G. Fratesi, S. de Gironcoli, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch. *J. of Phys. Condesn. Matter*, 21:395502, 2009.

[40] F. J. Giessibl. *Rev. Mod. Phys.*, 75:949, 2003.

[41] D. Goldfarb. *Math. Comp.*, 24:23, 1970.

[42] L. Gross, F. Mohn, N. Moll, P. Liljeroth, and G. Meyer. *Science*, 325:1110, 2009.

[43] L. Gross, F. Mohn, N. Moll, B. Schuler, A. Criado, E. Guitian, D. Pena, A. Gourdon, and G. Meyer. *Science*, 337:1326, 2012.

[44] L. Gross, B. Schuler, F. Mohn, N. Moll, N. Pavlicek, W. Steurer, I. Scivetti, K. Kotsis, M. Persson, and G. Meyer. *Phys. Rev. B*, 90:155455, 2014.

[45] C-S. Guo, M. A. Van Hove, X. Ren, and Y. Zhao. *J. Phys. Chem. C*, 119:1483, 2015.

[46] C-S. Guo, M. A. Van Hove, R-Q. Zhang, and C. Minot. *Langmuir*, 26:16271, 2010.

[47] C-S. Guo, X. Xin, M. A. Van Hove, X. Ren, and Y. Zhao. *J. Phys. Chem. C*, 119:14195, 2015.

[48] S. K. Hämäläinen, N. Heijden, J. Lit, S. Hartog, P. Liljeroth, and I. Swart. *Phys. Rev. Lett.*, 113:186102, 2014.

[49] D. R. Hamann, M. Schlüter, and C. Chiang. *Phys. Rev. Lett.*, 43:1494, 1979.

[50] J. Han, T.-L. Chan, and J. R. Chelikowsky. *Phys. Rev. B*, 82:153413, 2010.

[51] J. Han, M. L. Tiago, T.-L. Chan, and J. R. Chelikowsky. *J. Chem. Phys.*, 129:144109, 2008.

[52] P. Hapala, G. Kichin, C. Wagner, F. S. Tautz, R. Temirov, and P. Jelínek. *Phys. Rev. B*, 90:085421, 2014.

[53] P. Hapala, R. Temirov, F. S. Tautz, and P. Jelínek. *Phys. Rev. Lett.*, 113:226101, 2014.

[54] B. M. Haugerud, L. A. Bosworth, and R. E. Belford. *J. Appl. Phys.*, 94:4102, 2003.

[55] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136:B864, 1964.

[56] S. T. Huang and L. Yang. *App. Phys. Lett.*, 98:093114, 2011.

[57] X. Y. Huang, E. Lindgren, and J. R. Chelikowsky. *Phys. Rev. B*, 71:165328, 2005.

[58] K. H. Khoo, A. T. Zayak, H. Kwak, and J. R. Chelikowsky. First-principles study of confinement effects on the rahman spectra of si nanocrystals. *Phys. Rev. Lett.*, 105:115504, 2010.

[59] E. Kikuchi, S. Ishii, and K. Ohno. *Phys. Rev. B*, 74:195410, 2006.

[60] H. Kim and J. R. Chelikowsky. *Phys. Rev. Lett.*, 77:1063, 1996.

[61] M. Kim and J. R. Chelikowsky. *Appl. Surf. Sci.*, 303:163, 2014.

[62] M. Kim and J. R. Chelikowsky. *App. Phys. Lett.*, 107:163109, 2015.

[63] L. Kleinmann and D. M. Bylander. *Phys. Rev. Lett.*, 48:1425, 1982.

[64] W. Kohn, A. D. Becke, and R. G. Parr. *J. Phys. Chem.*, 100:12974, 1996.

[65] W. Kohn and L. J. Sham. *Phys. Rev.*, 140:A1133, 1965.

[66] A. Kokalj. *Comp. Mater. Sci.*, 28:155, 2003.

[67] Y. Kong, D. Shiri, and A. Buin. *Phys. Status Solidi RRL*, 3(9):281, 2009.

[68] J. Konle, H. Presting, H. Kibbel, K. Thonke, and R. Sauer. *Solid-State Elec.*, 45:1921, 2001.

[69] L. Kronik, A. Makmal, M. L. Tiago, M. M. G. Alemany, M. Jain, X. Y. Huang, Y. Saad, and J. R. Chelikowsky. *Phys. Stat. Sol.*, 243:1063, 2006.

[70] S. Lany and A. Zunger. *Phys. Rev. B*, 78:235104, 2008.

[71] A. J. Lee, T-L. Chan, and J. R. Chelikowsky. *Phys. Rev. B*, 89:075419, 2014.

[72] A. J. Lee, M. Kim, C. Lena, and J. R. Chelikowsky. *Phys. Rev. B*, 86:115331, 2012.

[73] A. J. Lee, Y. Sakai, M. Kim, and J. R. Chelikowsky. *App. Phys. Lett.*, 108:193102, 2016.

[74] M. L. Lee and E. A. Fitzgerald. *J. Appl. Phys.*, 94:2590, 2003.

[75] B. Li, A. F. Slachmuylders, B. Partoens, W. Magnus, and F. M. Peeters. *Phys. Rev. B*, 77:115335, 2008.

[76] L. Liu, M. Qing, Y. Wang, and S. Chen. *J. of Mater. Sci. and Tech.*, 31:599, 2015.

[77] P. Logan and X. H. Peng. *Phys. Rev. B*, 80:115322, 2009.

[78] S. G. Louie, S. Froyen, and M. L. Cohen. *Phys. Rev. B*, 26:1738, 1982.

[79] Y. Maeda, N. Tsukamoto, Y. Yazawa, Y. Kanemitsu, and Y. Masumoto. *Appl. Phys. Lett.*, 59:3168, 1991.

[80] G. Makov and M. C. Payne. *Phys. Rev. B*, 51:4014, 1995.

[81] J. B. McManus, R. People, R. L. Aggarwal, and P. A. Wolff. *J. Appl. Phys.*, 52:4748, 1981.

[82] D. Medaboina, V. Gade, S. K. R. Patil, and S. V. Khare. *Phys. Rev. B*, 76:205327, 2007.

[83] D. V. Melnikov and J. R. Chelikowsky. *Phys. Rev. B*, 69:113305, 2004.

[84] D. V. Melnikov and J. R. Chelikowsky. *Phys. Rev. Lett.*, 92:046802, 2004.

[85] F. Mohn, B. Schuler, L. Gross, and G. Meyer. *Appl. Phys. Lett.*, 102:073109, 2013.

[86] N. Moll, L. Gross, F. Mohn, A. Curioni, and G. Meyer. *New Journal of Physics*, 12:125020, 2010.

[87] K. Momma and F. Izumi. *J. Appl. Crystallogr.*, 44:1272, 2011.

[88] H. J. Monkhorst and J. D. Pack. *Phys. Rev. B*, 13:5188, 1976.

[89] F. D. Murnaghan. *Proc. Nat. Acad. Sci. USA*, 30:244, 1944.

[90] A. Natan, A. Benjamini, D. Naveh, L. Kronik, M. L. Tiago, S. P. Beckman, and J. R. Chelikowsky. *Phys. Rev. B*, 78:075109, 2008.

[91] G. Nesher, L. Kronik, and J. R. Chelikowsky. *Phys. Rev. B*, 71:035344, 2005.

[92] L. T. Ngo, D. Almécija, J. E. Sader, B. Daly, N. Petkov, J. D. Holmes, D. Erts, and J. J. Bolland. *Nano Lett.*, 6:2964, 2006.

[93] Y. M. Niquet, L. Genovese, C. Delerue, and T. Deutsch. *Phys. Rev. B*, 81:161301(R), 2010.

[94] R. Oberhuber, G. Zandler, and P. Vogl. *Phys. Rev. B*, 58:9941, 1998.

[95] M. Ondráček, P. Pou, V. Rozsíval, C. González, P. Jelínek, and R. Pérez. *Phys. Rev. Lett.*, 106:176101, 2011.

[96] S. Ossicini, M. Amato, R. Guerra, M. Palummo, and O. Pulci. *Nano. Res. Lett.*, 5:1637, 2010.

[97] N. Pavlicček, B. Fleury, M. Neu, J. Niedenführ, C. Herranz-Lancho, M. Ruben, and J. Repp. *Phys. Rev. Lett.*, 108:086101, 2012.

[98] N. Pavlicček, C. Herranz-Lancho, B. Fleury, M. Neu, J. Niedenführ, M. Ruben, and J. Repp. *Phys. Stat. Sol. B*, 250:2424, 2013.

[99] J. P. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.*, 77:3865, 1996.

[100] N. H. Quang, N. T. Truc, and Y. M. Niquet. *Comp. Mat. Sci.*, 44:21, 2008.

[101] R. Rurali, B. Aradi, T. Frauenheim, and Á. Gali. *Phys. Rev. B*, 79:115303, 2009.

[102] R. Rurali, M. Palummo, and X. Cartoixà. *Phys. Rev. B*, 81:235304, 2010.

[103] Y. Saad, Y. Zhou, C. Bekas, M. L. Tiago, and J. Chelikowsky. *Phys. Stat. Sol. B*, 243:2188, 2006.

[104] Y. Sakai, A. J. Lee, and J. R. Chelikowsky. *Nano. Lett.*, 16:3242, 2016.

[105] M. Schneiderbauer, M. Emmrich, A. J. Weymouth, and F. J. Giessibl. *Phys. Rev. Lett.*, 112:166102, 2014.

[106] D. F. Shanno. *M. Comp.*, 24:647, 1970.

[107] L. Shen, M. M. G. Alemany, and J. R. Chelikowsky. *J. Chem. Phys.*, 125:034311, 2006.

[108] H. S. Shin, J. Yu, J. Y. Song, and H. M. Park. *App. Phys. Lett.*, 94:011906, 2009.

[109] D. A. Smith, V. C. Holmberg, and B. A. Korgel. *ACS Nano*, 4:2356, 2010.

[110] A. Svizhenko, P. W. Leu, and K. Cho. *Phys. Rev. B*, 75:125417, 2007.

[111] A.M. Sweetman, S.P. Jarvis, H. Sang, I. Lekkas, P. Rahe, Y. Wang, J. Wang, N.R. Champness, L. Kantorovich, and P. Moriarty. *Nat. Comm.*, 5:3931, 2014.

[112] S. Z. Sze. *Physics of Semiconductor Devices*. Wiley, New York, 2 edition, 1981.

[113] K. Tanabe. *Energies*, 2:504, 2009.

[114] S. E. Taylor and F. Bruneval. *Phys. Rev. B*, 84:075155, 2011.

[115] M. Terrones, A. R. Botello-Méndez, J. Campos-Delgado, F. López-Urías, Y. I. Vega-Cantú, F. J. Rodríguez-Macías, A. L. Elías, E. Mu noz Sandoval, A. G. Cano-Márquez, J-C. Charlier, and H. Terrones. *Nano Today*, 5:351, 2010.

[116] B. Tian, X. Zheng, T. J. Kempa, Y. Fang, N. Yu, G. Yu, J. Huang, and C. M. Lieber. *Nature*, 449:885, 2007.

[117] A. Tkatchenko and M. Scheffler. *Phys. Rev. Lett.*, 102:073005, 2009.

[118] F. Tran and T. A. Wesolowski. *Int. J. Quantum Chem.*, 89:441, 2002.

[119] N. Troullier and J. L. Martins. *Phys. Rev. B*, 43:1993, 1991.

[120] D. Vanderbilt. *Phys. Rev. B*, 41:7892(R), 1990.

[121] L. Vicarelli, S. J. Heerema, C. Dekker, and H. W. Zandbergen. *ACS Nano*, 9:3428, 2015.

[122] W. Wang, R. S. Wu, and C. L. Yuan. *Phys. B-Con. Mat.*, 403:3997, 2008.

[123] A. J. Weymouth, T. Hofmann, and F. J. Giessibl. *Science*, 343:1120, 2014.

[124] A. J. Williamson, J. C. Grossman, R. Q. Hood, A. Puzder, and G. Galli. *Phys. Rev. Lett.*, 89:196803, 2002.

[125] J. J. Wortman and R. A. Evans. *J. Appl. Phys.*, 36:153, 1965.

[126] Y. Wu, Y. Cui, L. Huynh, C. J. Barrelet, D. C. Bell, and C. M. Lieber. *Nano Lett.*, 4:433, 2004.

[127] J. Yoon, A. M. Girgis, I. Shalish, L. R. Ram-Mohan, and V. Narayanamurti. *Appl. Phys. Lett.*, 94:142102, 2009.

[128] C. Zhang, A. De Sarkar, and R. Q. Zhang. *J. Phys-Con. Mat.*, 24:015301, 2012.

[129] F. Zhang and V. H. Crespi. *Phys. Rev. Lett.*, 102:156401, 2009.

[130] J. Zhang, P. Chen, B. Yuan, W. Ji, Z. Cheng, and X. Qiu. *Science*, 342:611, 2013.

[131] X. Zhao, C. M. Wei, L. Yang, and M. Y. Chou. *Phys. Rev. Lett.*, 92:236805, 2004.

[132] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky. *Phys. Rev. E*, 74:066704, 2006.

[133] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky. *J. of Comp. Phys.*, 219:172, 2006.